

Abuse and dependence on prescription opioids in adults: a mixture categorical and dimensional approach to diagnostic classification

L.-T. Wu^{1*}, G. E. Woody², C. Yang³, J.-J. Pan⁴ and D. G. Blazer¹

¹ Department of Psychiatry and Behavioral Sciences, Duke University School of Medicine, Duke Clinical Research Institute, Durham, NC, USA

² Department of Psychiatry, School of Medicine, University of Pennsylvania and Treatment Research Institute, Philadelphia, PA, USA

³ Social Science Research Institute, Duke University, Durham, NC, USA

⁴ Veterans Health Administration, Washington, DC, USA

Background. For the emerging DSM-V, it has been recommended that dimensional and categorical methods be used simultaneously in diagnostic classification; however, little is known about this combined approach for abuse and dependence.

Method. Using data ($n=37\,708$) from the 2007 National Survey on Drug Use and Health (NSDUH), DSM-IV criteria for prescription opioid abuse and dependence among non-prescribed opioid users ($n=3037$) were examined using factor analysis (FA), latent class analysis (LCA, categorical), item response theory (IRT, dimensional), and factor mixture (hybrid) approaches.

Results. A two-class factor mixture model (FMM) combining features of categorical latent classes and dimensional IRT estimates empirically fitted more parsimoniously to abuse and dependence criteria data than models from FA, LCA and IRT procedures respectively. This mixture model included a severely affected group (7%) with a comparatively moderate to high probability (0.32–0.88) of endorsing all abuse and dependence criteria items, and a less severely affected group (93%) with a low probability (0.003–0.16) of endorsing all criteria. The two empirically defined groups differed significantly in the pattern of non-prescribed opioid use, co-morbid major depression, and substance abuse treatment use.

Conclusions. A factor mixture model integrating categorical and dimensional features of classification fits better to DSM-IV criteria for prescription opioid abuse and dependence in adults than a categorical or dimensional approach. Research is needed to examine the utility of this mixture classification for substance use disorders and treatment response.

Received 5 January 2010; Revised 30 March 2010; Accepted 3 April 2010; First published online 12 May 2010

Key words: DSM-IV, factor mixture model, item response theory, latent class analyses, nosology, prescription opioid use disorders.

Introduction

Indicators from multiple sources show a substantial increase in non-prescribed opioid use (i.e. non-prescribed use of prescription opioids), in addition to prescription opioid-related treatment admissions and mortality (Zacny *et al.* 2003; Manchikanti, 2007; SAMHSA, 2009). Recent data show an estimated 2.2 million new, non-prescribed prescription opioid users/year, with most new users (62%) aged ≥ 18 years (SAMHSA, 2009). In 2008, 1.7 million Americans

had a current prescription opioid use disorder (OUD) according to DSM-IV criteria, representing 79% of those with any past-year prescription drug use disorder (SAMHSA, 2009). After cannabinoids, prescription OUDs are the second most prevalent drug use disorder in the USA (SAMHSA, 2009).

DSM-V is scheduled to be published in 2013 (Kupfer *et al.* 2008; Schatzberg, 2010), yet empirical data informing diagnostic classification of current prescription OUDs in adults are lacking. DSM-IV's categorical distinction between dependence and abuse has been controversial because of mixed findings (Teesson *et al.* 2002; Proudfoot *et al.* 2006; Gillespie *et al.* 2007; Helzer *et al.* 2007; Lynskey & Agrawal, 2007; Wu *et al.* 2009c). A dimensional approach taking into account the variability of a measured condition

* Address for correspondence: L.-T. Wu, Sc.D., Department of Psychiatry and Behavioral Sciences, Duke University School of Medicine, Duke Clinical Research Institute, Duke University Medical Center, Durham, NC 27710, USA.
(Email: litzy.wu@duke.edu)

has been considered for DSM-V to complement the categorical approach (Helzer *et al.* 2007); however, existing research relies mainly on either a categorical or a dimensional approach, making the issue of using both approaches simultaneously difficult to resolve. We address this gap by applying categorical, dimensional and mixture models to evaluate the dimensionality of DSM-IV criteria as they apply to prescription OUDs.

Earlier studies using factor analysis (FA) to inform categories found that DSM-III-R and DSM-IV dependence criteria represent one factor; however, they often relied on a small sample of treatment-seeking opioid users, examined dependence symptoms only, and did not consider criterion-level psychometric performance (Kosten *et al.* 1987; Morgenstern *et al.* 1994; Feingold & Rounsaville, 1995). Subsequently, Nelson *et al.* (1999) used FA to examine DSM-IV's criteria for alcohol, cannabis and opioid users. They found that a one-factor model explained the criteria for each disorder, but when the analysis was restricted to the subsample with low to moderate symptoms, a two-factor model (dependence *versus* abuse) was noted. Beseler *et al.* (2006) applied FA to examine criteria of drug use disorders in a sample of men from the Virginia Twin Registry and found that a two-factor model (physical dependence *versus* adverse consequences) explained OUD symptoms.

Recently, we and other investigators turned to a dimensional approach using item response theory (IRT) analysis to evaluate the dimensionality and criterion-specific performance of DSM-IV criteria for drug use disorders (Wu *et al.* 2009*b,c*). For example, results from IRT analyses of DSM-IV criteria for specific drug use disorders in a sample of men from the Virginia Twin Registry (Gillespie *et al.* 2007) and respondents to the National Epidemiologic Survey on Alcohol and Related Conditions (NESARC) (Lynskey & Agrawal, 2007), and of DSM-IV criteria for alcohol and marijuana use disorders in an Australian sample (Teesson *et al.* 2002; Proudfoot *et al.* 2006) showed that the abuse and dependence criteria for each substance class reflected a single underlying continuum. Additionally, the 'problems with the law' criterion seemed to tap the extreme form of substance use problems, suggesting that the DSM-IV's categorical distinction between dependence and abuse may not be justified.

Although IRT analysis describes a dimensional construct of a condition, it does not produce diagnostic categories (Muthén & Asparouhov, 2006). Such categories, nonetheless, are important because they facilitate communications among individuals, comparisons of health statistics, and insurance payment for treatments (APA, 2000). Additionally, if diagnostic

categories are continuous in nature, a mixture model that considers both categories and dimensions is expected to fit diagnostic data better than an IRT model. A newer factor mixture model (FMM), which combines features of dimensional (IRT or FA) and categorical (latent class analysis; LCA) analyses into a hybrid model, has promise for evaluating the dimensional *versus* categorical nature of DSM-IV criteria (Muthén, 2006; Muthén & Asparouhov, 2006). LCA has been used to identify subtypes (categories) of substance dependence (Bucholz *et al.* 1996; Wu *et al.* 2009*a*), but it does not address the continuous nature of a condition. Notably, an FMM not only generates categories of individuals but also addresses this limitation of a categorical approach by including a continuous component to describe the variability of a condition. An FMM provides both empirically defined categories (latent classes) and a continuous factor (IRT estimates) within each class. Muthén & Asparouhov (2006) have illustrated the application of an FMM to diagnostic criteria for tobacco dependence and alcohol use disorders (Muthén, 2006). Kuo *et al.* (2008) applied it to diagnostic criteria for alcohol dependence and found that data of alcohol dependence criteria fit well to an FMM. However, the FMM has not been applied to evaluate criteria for drug abuse and dependence. Here we apply the FMM (in addition to FA, IRT and LCA procedures) to evaluate the dimensionality and criterion-specific performance of DSM-IV criteria for prescription OUDs. We also explore potential external validators of diagnostic categories (patterns of non-prescribed opioid use, co-morbid major depression, use of substance abuse treatment) to characterize empirically defined groups (Wu *et al.* 2008*b*).

We address the following questions: (1) does an FMM empirically describe OUD criteria better than FA, LCA and IRT models; and (2) if so, are FMM-defined categories associated with distinct demographics (gender, age, race/ethnicity, educational level, and family income), patterns of non-prescribed opioid use (age of first use and frequency of use), and severity-related factors (co-morbid major depression and substance abuse treatment use)? Based on recent evidence that abuse and dependence may not represent two diagnostic groups (Proudfoot *et al.* 2006; Gillespie *et al.* 2007; Lynskey & Agrawal, 2007), we hypothesized the presence of one affected (high-risk) group and another less affected (low-risk) group, in contrast to three groups as defined by DSM-IV (dependence, abuse, none). Consistent with IRT findings that substance use problems are thought to be continuous, individuals within each group were hypothesized to have varying degrees of opioid-related problems. A two-class FMM will provide a better fit to the OUD criteria than an FA, IRT or LCA model

because of the FMM's integration of categorical and dimensional features. The affected group will be more likely than the less affected group to exhibit a pattern of frequent non-prescribed opioid use, have more major depression, and use more substance abuse treatment.

Method

Data source

Data were from the public-use data file of the 2007 National Survey on Drug Use and Health (NSDUH), the largest sample of adult non-prescribed opioid users and the only survey designed to provide ongoing national estimates of substance use and disorders in the USA (SAMHSA, 2008). The survey uses multi-stage area probability sampling methods to select a representative sample of the US civilian, non-institutionalized population aged ≥ 12 years. Participants include household residents; residents of shelters, rooming/boarding houses, half-way houses, college dormitories, and group homes; and civilians residing on military bases.

Prospective respondents are assured that their names will not be recorded and their responses will be kept strictly confidential, and all study procedures and protections are carefully explained. The survey combines computer-assisted personal interviewing and audio computer-assisted self-interviewing (ACASI) to increase honest reporting of substance use and related behaviors (Turner *et al.* 1998). ACASI provides the respondent with a confidential means of responding and is used for questions of a sensitive nature (substance use and mental health). Respondents read questions on the computer screen or questions are read to respondents through headphones; they enter their responses directly into a laptop computer provided by the interviewer.

In 2007, a total of 67870 respondents aged ≥ 12 years completed the interview with a weighted interviewing response rate of 74%. NSDUH uses methods to ensure that its sample is representative of the general population aged ≥ 12 years (SAMHSA, 2008).

This study was exempt from review by the Duke Institutional Review Board because the data are available in the public domain without any identification of personal information.

Study variables

Non-prescribed (prescription) opioid use was defined as any self-reported use of prescription opioid pain relievers that were not prescribed for the respondent or that the respondent took only for the experience or

feeling they caused (Wu *et al.* 2008*a,b*). Respondents were read the following statement: 'These questions are about prescription pain reliever use. We are not interested in your use of "over-the-counter" pain relievers such as aspirin, Tylenol[®], or Advil[®], which can be bought in drug stores/grocery stores without a doctor's prescription.' The survey then used discrete questions and presented a card showing pictures of prescription pain relievers (e.g. Vicodin[®], Lortab[®], Darvocet[®], codeine, Percocet[®], Tylox[®], hydrocodone, and OxyContin[®]) to assess respondents' non-prescribed use, age of first non-prescribed use, and the number of days of using any non-prescribed opioids within the past 12 months.

Assessments for prescription OUDs were specified by DSM-IV criteria (APA, 2000). The four dichotomous abuse criteria include: (A1) serious problems at home, work, or school; (A2) regular consumption that put the user in physical danger; (A3) repeated use that led to trouble with the law; and (A4) problems with family or friends caused by continued use. The seven dichotomous dependence criteria are: (D1) tolerance; (D2) withdrawal; (D3) more frequent use than intended or inability to maintain limits on use; (D4) inability to reduce or stop use; (D5) spending a great deal of time over a period of a month using the drugs or getting over the effects of use; (D6) reduced involvement or participation in important activities because of use; and (D7) continued use despite related problems with emotions, nerves, mental or physical health.

We examined whether empirically defined groups of non-prescribed users differed in age, gender, race/ethnicity, educational level, and annual family income to elucidate their key demographic profiles. To explore the utility of empirically defined groups with regard to their association with other clinical indicators, we examined whether the affected group was more likely than the less affected group to use substance abuse treatment and have co-morbid major depression (Wu *et al.* 2008*b*). Substance abuse treatment use was defined as any receipt of treatment services specifically related to alcohol or drug use in the prior year. Past-year DSM-IV major depressive episodes were assessed by questions adapted from National Comorbidity Survey Replication (SAMHSA, 2008). Other mental disorders were not assessed.

Data analysis

SUDAAN (2006) was used to examine the distributions of study variables. The 11 criteria (four abuse and seven dependence criteria) were analyzed by Mplus using FA, IRT, LCA and FMM procedures (Muthén & Muthén, 2007). As research suggests a

one-factor construct underlying abuse and dependence criteria, in contrast to two discrete conditions (Gillespie *et al.* 2007; Lynskey & Agrawal, 2007), we report factor loadings from a one-factor model (abuse and dependence criteria) and a two-factor model (abuse criteria *versus* dependence criteria). The Tucker–Lewis Index (TLI), Comparative Fit Index (CFI) and root mean square error of approximation (RMSEA) were used to assess the model fit of FA. Values of TLI and CFI ≥ 0.95 (1 = perfect fit) and values of RMSEA ≤ 0.06 indicate an excellent fit to the data (Browne & Cudeck, 1993; Hu & Bentler, 1999).

Consistent with prior IRT research, we used a two-parameter IRT model to examine the latent continuum of the 11 OUD criteria, which assumes that a one-factor model provides the most parsimonious fit to the data (Gillespie *et al.* 2007; Lynskey & Agrawal, 2007; Wu *et al.* 2009c). Two-parameter logistic IRT modeling was conducted to determine the relationship between respondents' item performance and the latent severity (trait) of opioid use problems, which is described by a monotonically increasing S-shaped item characteristic curve (ICC) (Wu *et al.* 2009b,c). An ICC is characterized by item severity and discrimination parameters. A severity parameter indicates the position of the ICC in relation to the latent continuum (typically -3 to $+3$). It describes the severity level best measured by a specific item, and reflects the point on the latent continuum where there is a 50% chance of the criterion being present. A discrimination parameter measures the precision with which the item distinguishes between respondents with levels of the latent trait above *versus* those with levels below the item's severity (Langenbucher *et al.* 2004). Items with high discrimination values (steep slopes) are more useful for discriminating between opioid users above or below given levels on the continuum than are items with low values.

Next, we applied the 11 criteria to LCA to determine different latent classes/categories. A two-class model was estimated first. The Vuong–Lo–Mendell–Rubin likelihood ratio (VLMR) test showed significant differences between one-class and two-class models ($p < 0.001$), suggesting that a two-class model had a better fit (Lo *et al.* 2001). Additionally, there were no differences between two-class and three-class models (VLMR test; $p > 0.05$), or between three-class and four-class models (VLMR test; $p > 0.05$). Together, they suggested that a two-class model was acceptable within the context of LCA.

The 11 criteria were then analyzed by FMM (Muthén & Asparouhov, 2006). FMM offers the possibility to examine models that contain both categorical classes and continuous factors within each class by allowing for the presence of between-class

heterogeneity and within-class variation. Specifically, whereas LCA (categorical latent variables) identifies homogeneous groups of individuals, LCA-defined categories ignore the possible within-class heterogeneity in a form of variation of severity. IRT (continuous latent variables) describes underlying dimensions of diagnostic criteria; however, there are no model-based categories and natural thresholds for diagnosis are difficult to specify (Muthén, 2006). FMM addresses their limitations by providing a bridge between categorical and continuous latent approaches through incorporating within-category variation in a form of one or more continuous factors (IRT estimates for one factor; FA estimates for two or more factors) into categorical latent variable (LCA) modeling. Under FMM, non-prescribed opioid users are characterized by categories of liability, but liability within each category is distributed continuously (latent traits). Based on DSM-IV and research findings (APA, 2000; Wu *et al.* 2009c), we compared two-class with three-class FMMs. Because abuse and dependence criteria are highly correlated (Teesson *et al.* 2002; Proudfoot *et al.* 2006; Wu *et al.* 2009c), we hypothesized the presence of one factor (unidimensionality) within each class. In this context, it is an IRT mixture model (Muthén & Asparouhov, 2006).

Furthermore, non-nested models of FA, IRT, LCA and FMM were compared by considering the number of parameters, Bayesian information criterion (BIC), and sample-size adjusted BIC (ABIC) (Muthén, 2006). The VLMR test is not used for comparing non-nested models. BIC/ABIC are reported to outperform the Akaike information criterion (AIC) in evaluating a fit to observed responses (Nylund *et al.* 2007). BIC/ABIC takes into account the parsimony of the model (number of parameters); lower values indicate a better model. In addition, entropy (range 0–1) was considered for evaluating how well latent classes can be distinguished, with higher values indicating a clearer delineation of classes (Nylund *et al.* 2007). Although entropy is not a fit indicator for FA and IRT models (not used for comparing non-nested models), it provides additional information about the quality of classes from LCA or FMM.

Finally, we examined whether empirically defined groups of non-prescribed users differed in demographics, patterns of non-prescribed opioid use, substance abuse treatment use, and major depression using χ^2 and logistic regression procedures in SUDAAN to take into account weighting and clustering of the data. They determined demographic disparities in problematic opioid use across subgroups, explored their association with potential needs for treatment, and identified at-risk groups for future research. To be representative of the NHDUH sample,

Table 1. Prevalence of opioid abuse and dependence symptoms and their factor loadings^a among past-year non-prescribed users of prescription opioids aged ≥ 18 years ($n = 3037$)

DSM-IV criteria	Prevalence % (s.e.)	One-factor model ^b (abuse/dependence)	Two-factor model ^c (abuse <i>v.</i> dependence)	
A1: Role interference	5.5 (0.92)	0.96	0.97	–
A2: Hazardous use	6.8 (0.85)	0.85	0.86	–
A3: Problems with the law	2.5 (0.52)	0.85	0.86	–
A4: Relationship problems	4.5 (0.83)	0.97	0.98	–
D1: Tolerance	19.8 (1.32)	0.84	–	0.84
D2: Withdrawal	8.1 (0.85)	0.80	–	0.80
D3: Taking larger amounts	4.3 (0.58)	0.86	–	0.86
D4: Inability to cut down	4.5 (0.71)	0.84	–	0.84
D5: Time spent using	16.1 (1.44)	0.82	–	0.82
D6: Giving up activities	7.6 (1.01)	0.97	–	0.98
D7: Continued use despite resulting problems	7.5 (0.97)	0.91	–	0.92

s.e., Standard error.

^a Factor loadings were freely estimated.

^b Comparative Fit Index (CFI)=0.99; Tucker–Lewis Index (TLI)=0.99; root mean square error of approximation (RMSEA)=0.03.

^c CFI=0.99; TLI=0.99; RMSEA=0.03.

data were weighted to adjust for variation in household selection, non-response and post-stratification of the sample characteristics to census data. All results are weighted estimates taking into account weighting and clustering, except for sample sizes (unweighted).

Results

Demographics of non-prescribed opioid users

Of 37708 adults aged ≥ 18 years, 5% ($n = 3037$) used non-prescribed prescription opioids in the past year. Of these 3037 users, 43% were females, 59% were aged 18–34 years, 26% were non-whites (12% blacks, 5% Hispanics, 9% others), 52% had not attended a college, and 49% had a family income under US\$40000.

Prevalence of opioid abuse and dependence criteria

Of the 11 criteria, ‘tolerance’ (D1, 20% among all users) and ‘time spent’ (D5, 16%) were most often endorsed, whereas ‘problems with the law’ (A3, 2.5%) had a much lower prevalence. Overall, 10.3% of these 3037 users met criteria for opioid dependence (≥ 3 dependence criteria), and another 3.5% met criteria for opioid abuse (≥ 1 abuse criterion and absence of a dependence diagnosis). ‘Hazardous use’ (A2, 69%) was the most commonly endorsed symptom in the abuse group; for those in the dependence group, ‘tolerance’ (D1, 89%), ‘time spent’ (D5, 81%), ‘continued use’ (D7, 65%), ‘withdrawal’ (D2, 63%), and ‘giving

up activities’ (D6, 61%) were endorsed by more than one half.

Factor loadings of abuse and dependence criteria (Table 1)

The results from either a one-factor model (abuse/dependence) or a two-factor model (abuse *versus* dependence) showed a good to high level of factor loadings (≥ 0.80 for each item). The two-factor model indicated that the four abuse criteria and the seven dependence criteria were highly correlated (0.98).

Model comparisons from FA, IRT, LCA and FMM (Table 2)

As shown in Table 2, two-class and three-class FMMs had lower BIC/ABIC (more parsimonious) than the other models. The VLMR test comparing two-class with three-class FMMs indicated no differences in model fit ($p > 0.05$). To provide a less biased and more parsimonious understanding of the latent abuse/dependence trait across groups, a two-class FMM was estimated with factor loadings constrained to be equal across groups (to reduce measurement errors in between-group comparisons) and thresholds (severity estimates) to be freely estimated. This constrained FMM allowed individual discrimination parameters within class but variant severity parameters within and across classes. This constrained FMM had a high quality of classification of classes (entropy = 0.96; very low classification errors).

Table 2. Comparisons of factor analysis (FA), item response theory (IRT), latent class analysis (LCA) and factor mixture model (FMM) analyses of DSM-IV criteria for prescription opioid abuse and dependence among past-year non-prescribed users of prescription opioids aged ≥18 years (n = 3037)

Model	No. of parameters	Log L	ABIC	BIC	Entropy
One-factor (IRT) model	22	-5918.38	11943.27	12013.2	N.A.
Two-factor model abuse v. dependence	23	-5910.03	11931.41	12004.5	N.A.
LCA					
Two-class	23	-6200.97	12513.30	12586.4	0.96
Three-class	35	-5864.61	11898.65	12009.9	0.92
Four-class	47	-5785.35	11798.23	11947.6	0.94
FMM: two-class one-factor					
Unconstrained	47	-5762.14	11751.82	11901.16	0.72
Equal factor loadings ^a	36	-5802.77	11779.82	11894.21	0.96
FMM: three-class one-factor					
Unconstrained	72	-5672.24	11693.05	11921.82	0.86
Equal factor loadings ^a	50	-5761.29	11764.65	11923.52	0.61

ABIC, Adjusted Bayesian information criterion; N.A., not available.

Models of IRT, LCA and FMM were obtained using the maximum likelihood estimator.

^a With factor loadings constrained to be equal across groups and thresholds to be freely estimated.

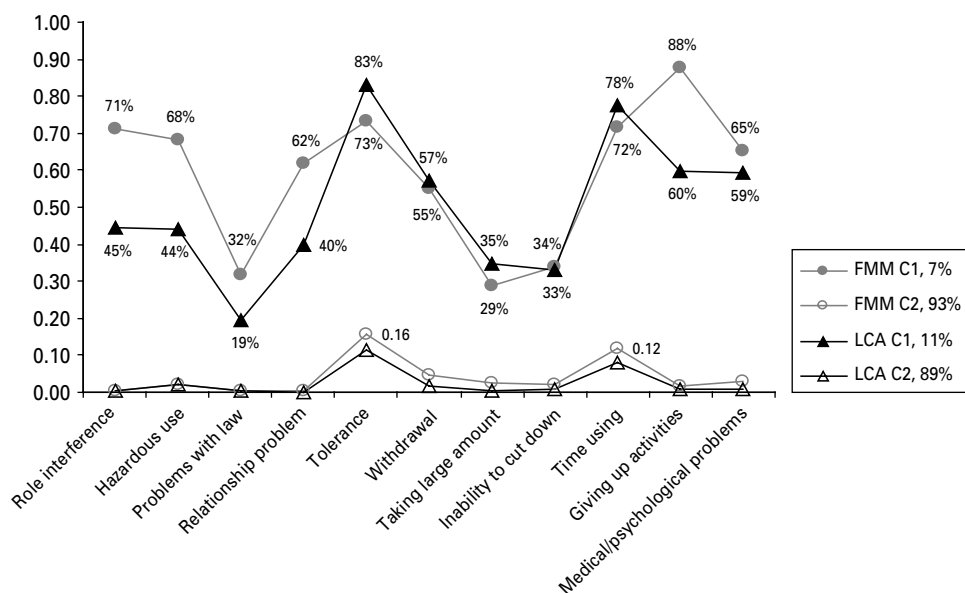


Fig. 1. Symptom profiles of a two-class factor mixture model (FMM) and a two-class latent class analysis (LCA) of prescription opioid abuse and dependence criteria among past-year non-prescribed users of prescription opioids aged ≥18 years (n = 3037).

FMM-defined diagnostic groups (Fig. 1)

The constrained two-class FMM included an affected group (7%) and a less affected group (93%). The affected group had a higher probability of endorsing the 11 criteria (0.29–0.88) than the less affected group (0.003–0.16). Their symptom profiles were similar to those from a two-class LCA (Fig. 1).

ICCs: item discrimination and severity (Table 3; Fig. 2)

Item-level discrimination and severity parameters from the constrained two-class FMM (Table 3) that describe the relationship between opioid users’ item performance and the latent trait underlying OUD criteria (ICCs) are shown in Fig. 2. The 11 criteria had

Table 3. Item discrimination and item severity estimates from a two-class factor mixture model (FMM)^a of past-year non-prescribed users of prescription opioids aged ≥ 18 years ($n = 3037$)

DSM-IV criteria ^a	Class 1: Affected group		Class 2: Less affected group	
	Item discrimination	Item severity	Item discrimination	Item severity
A1: Role interference	0.51 (0.24)	-1.21 (0.67)	0.94 (0.34)	4.05 (0.96)
A2: Hazardous use	0.24 (0.09)	-1.97 (1.11)	0.44 (0.14)	5.44 (1.38)
A3: Problems with the law	0.29 (0.15)	1.67 (0.93)	0.52 (0.27)	6.88 (2.98)
A4: Relationship problems	1.07 (0.28)	-0.42 (0.26)	1.97 (0.46)	3.21 (0.48)
D1: Tolerance	1.00 (0.23)	-0.89 (0.33)	1.82 (0.34)	1.17 (0.21)
D2: Withdrawal	0.64 (0.13)	-0.24 (0.29)	1.17 (0.15)	2.28 (0.08)
D3: Taking larger amounts	1.41 (0.34)	0.69 (0.23)	2.58 (0.52)	2.14 (0.18)
D4: Inability to cut down	0.78 (0.21)	0.67 (0.27)	1.43 (0.30)	2.50 (0.11)
D5: Time spent using	0.72 (0.16)	-1.00 (0.37)	1.33 (0.22)	1.51 (0.10)
D6: Giving up activities	0.59 (0.16)	-2.32 (0.84)	1.08 (0.19)	3.08 (0.26)
D7: Continued use despite resulting problems	0.95 (0.22)	-0.57 (0.24)	1.74 (0.23)	2.19 (0.12)

Standard error (S.E.) given in parentheses.

^a With factor loadings constrained to be equal across groups and thresholds to be freely estimated.

Goodness-of-fit indices [the Comparative Fit Index (CFI), Tucker-Lewis Index (TLI) and the root mean square error of approximation (RMSEA)] are not available in a mixture model.

low to high discrimination parameters (0.24–1.41) in distinguishing between opioid users and tapped a low to high severe range (-2.32 to 1.67) of the affected group. The criteria in the less affected group generally had higher discrimination values (0.44–2.58) in distinguishing between opioid users and to measure severe ranges of opioid use problems (1.17–6.88). In both groups, A2 (hazardous use) and A3 (problems with the law) had lower discrimination values (less reliable) than other items in distinguishing between opioid users [flatter curves in Fig. 2(a, b)]. Because of the low frequencies of abuse/dependence symptoms in the less affected group, the standard errors of the estimates in this group were somewhat larger. Both groups also differed in item-level severity. A3, D3 (taking larger amounts) and D4 (inability to cut down) tapped the most severe ranges of the affected group, whereas all abuse criteria measured the most severe end of the less affected group (shifted to the right side in Fig. 2b).

Exploratory external validators (Table 4)

T-tests showed that, relative to the less affected group, the affected group had a higher number of OUD criteria met (mean: 6.76 *v.* 0.46, $p < 0.001$) and greater number of days of using non-prescribed opioids in the past year (mean: 132.92 *v.* 38.8, $p < 0.001$).

Finally, χ^2 and logistic regression procedures examined the association of FMM-defined groups with age, gender, race/ethnicity, education, family income,

major depression, substance abuse treatment use, and the frequency and age of first non-prescribed opioid use (Table 4). Compared with the less affected group, a high-school education [adjusted odds ratio (aOR) 1.89], major depression (aOR 2.91), use of substance abuse treatment (aOR 5.99) and non-prescribed opioid use on ≥ 12 days in the past year (aOR 5.35) were associated with being in the affected group.

Discussion

Using data from a nationally representative sample of non-prescribed prescription opioid users, this research is the first to compare categorical, dimensional and mixture approaches in examining the latent construct of DSM-IV criteria for OUDs. Our study findings not only add new information to classification of OUDs but also serve as a stepping stone for further investigation of this approach to diagnostic criteria for other substances. An FMM incorporating categorical LCA and dimensional IRT components empirically fitted more parsimoniously to OUD criteria than LCA or IRT models. The results from the FMM suggested the presence of two groups of non-prescribed opioid users, and the severity of opioid use problems within each group could be conceptualized as continuous. The two FMM-defined groups differed in the pattern of non-prescribed opioid use, co-morbid major depression, and substance abuse treatment use. Therefore, a categorical approach to diagnosis may be

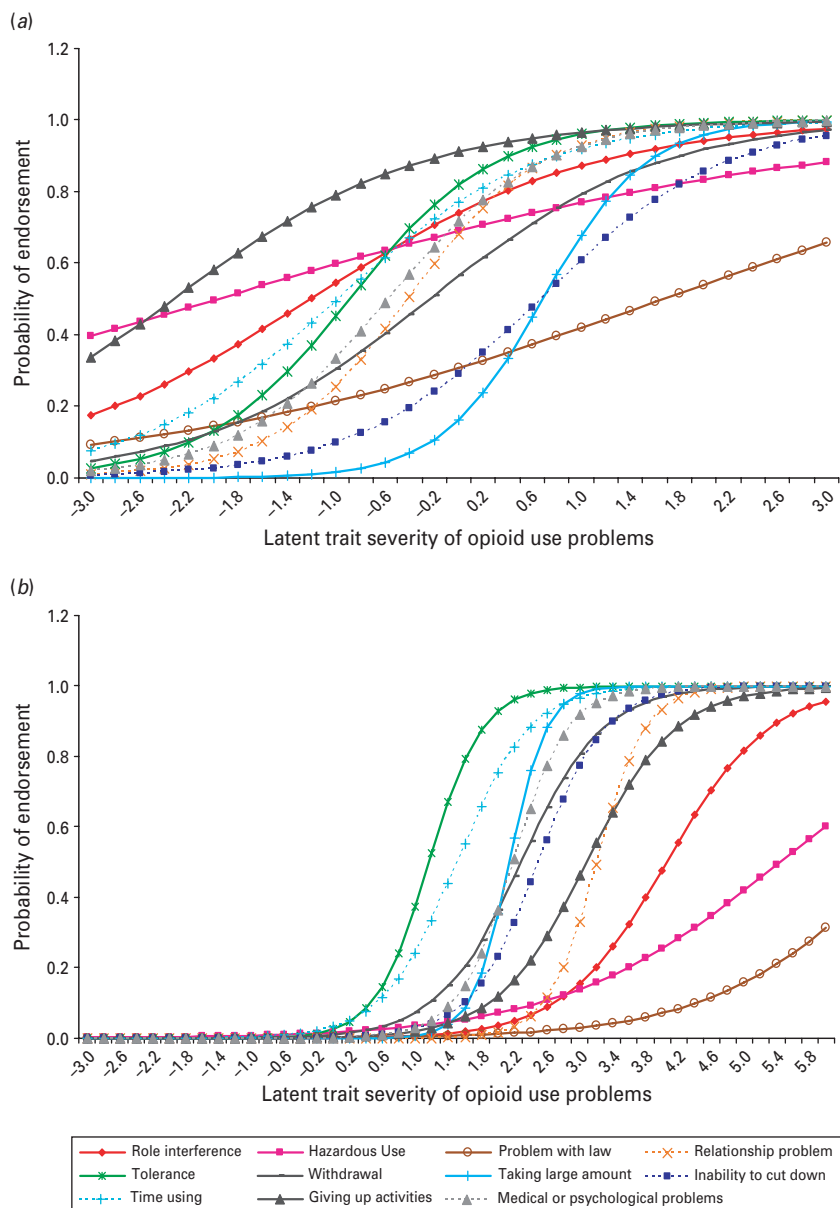


Fig. 2. Item characteristic curves (ICCs) for the prescription opioid abuse and dependence criteria: (a) class 1: the severely affected group; (b) class 2: the less affected group. An ICC is characterized by item severity and discrimination parameters. It relates the probability of endorsing criterion symptoms to opioid use problems.

enhanced by adding continuous features to better characterize between-group heterogeneity in the risk and within-group variability in manifestations of symptoms.

A mixture of categorical and dimensional constructs

Notably, although FA and IRT procedures indicated a unidimensional construct underlying OUD criteria, the results from LCA and FMM revealed further the presence of heterogeneity (different classes) underlying them. Empirically, FMM fitted OUD criteria

more parsimoniously than a pure dimensional or categorical model. Recent FMM analyses of alcohol and tobacco dependence criteria also found that an FMM explained diagnostic criteria more parsimoniously than an LCA or IRT model (Muthén & Asparouhov, 2006; Kuo et al. 2008). It is likely that an FMM improves the model fit because it allows for the presence of different groups within a sample by taking into account heterogeneity in the manifestations of OUD symptoms while simultaneously recognizing the continuous nature of measured conditions (IRT estimates). This finding is in line with results from LCA

Table 4. A two-class factor mixture model (FMM)^a: adjusted odds ratios (aORs) and 95% confidence intervals (CIs) of FMM-defined groups of adult non-prescribed users of prescription opioids aged ≥ 18 years ($n = 3037$)

Selected characteristics	Class 1: Affected (%)	Class 2: Less affected (%)	χ^2 (df) <i>p</i> value	Affected group <i>v.</i> less affected group ^b aOR (95% CI)
Gender				
Male	61.1	56.5	0.49 (1)	1.00 (0.60–1.67)
Female	38.9	43.5	n.s.	1.00
Age group				
18–25 years	29.5	38.0		0.61 (0.28–1.32)
26–34 years	20.8	21.7	2.07 (2)	0.82 (0.38–1.74)
≥ 35 years	49.7	40.2	n.s.	1.00
Race/ethnicity				
White	63.3	75.1		1.00
Black	17.9	8.5		1.33 (0.46–3.84)
Hispanic	14.5	11.5	2.31 (3)	1.52 (0.57–4.07)
Other	4.2	4.9	n.s.	1.06 (0.42–2.70)
Education				
<High school	31.6	19.2		1.86 (0.94–3.68)
High school	43.9	31.8	15.76 (2)	1.89 (1.14–3.14)
\geq College	24.5	49.1	<0.01	1.00
Family income				
<US\$40000	65.9	47.9		1.18 (0.68–2.05)
US\$40000–US\$74999	17.7	28.4	8.80 (2)	0.79 (0.41–1.53)
\geq US\$75000	16.4	23.8	<0.01	1.00
Past-year major depression				
Yes	42.2	17.0	7.83 (1)	2.91 (1.73–4.89)
No	57.9	83.0	<0.01	1.00
Past-year use of substance abuse treatment				
Yes	37.3	5.9	10.54 (1)	5.99 (3.00–11.96)
No	62.7	94.0	<0.01	1.00
Total number of days using non-prescribed opioids in the past year				
≥ 12 days	86.7	48.2	30.22 (1)	5.35 (2.72–10.52)
1–11 days	13.3	51.8	<0.01	1.00
Age of first non-prescribed opioid use				
≤ 17 years	62.3	70.1	2.14 (1)	0.64 (0.36–1.16)
≥ 18 years	37.7	29.9	n.s.	1.00

df, Degrees of freedom; n.s., not significant ($p > 0.05$). Bold type indicates $p < 0.05$.

^a With factor loadings constrained to be equal across groups and thresholds to be freely estimated.

^b The adjusted logistic regression model included all variables listed in the first column.

and IRT research. For example, studies of substance dependence suggest that LCA-defined subtypes are graded in severity (Bucholz *et al.* 1996; Wu *et al.* 2009a). Similarly, IRT analyses of DSM-IV criteria for drug use diagnoses in adults (Gillespie *et al.* 2007; Lynskey & Agrawal, 2007) and for prescription OUDs in adolescents (Wu *et al.* 2009c) found that abuse/dependence criteria were a unidimensional construct.

However, comparisons between studies are complicated because a single classification approach is

applied. Instead, this research represents an advance by applying dimensional, categorical and mixture approaches to help understand the classification for DSM-IV criteria. The results from the mixture approach suggest the presence of two groups of non-prescribed opioid users characterized by high and low opioid use problems respectively, and opioid use problems within each group tend to vary continuously. They suggest some support for a categorical approach to identifying different risk groups of drug

users for tailored interventions or clinical research (e.g. motivational interviews for those with low problems; more intensive treatments for those with high problems), and add evidence to prior IRT findings that drug use problems are dimensional conditions. Together, they imply a mixture of categorical and dimensional constructs for opioid use problems and demonstrate a need for further research on this conceptualization (e.g. FMM analyses of epidemiological samples of substance users; genetic research on risk and clinical markers across groups). This line of research is timely, given that categorical classification is proposed to continue in DSM-V, but the role of dimensional components to categorical diagnoses remains unclear.

Abuse criteria measuring comparatively severe symptoms

Another salient finding concerns variations in criterion-level performance within and across groups, suggesting the heterogeneity in the manifestation and perhaps liability to OUD problems across groups. For the affected group, item severity estimates for abuse and dependence criteria overlap, whereas all abuse criteria seem to measure a more severe subset of the less affected group than dependence criteria. This result supports prior IRT studies that dependence criteria are not necessarily more severe than abuse criteria and the DSM-IV's hierarchical distinction between dependence and abuse is not fully supported (Gillespie *et al.* 2007; Wu *et al.* 2009c). They also provide the new finding that abuse symptoms may be useful for identifying a problematic subset of opioid users. One possible area for future research is to examine whether drug users who develop abuse symptoms initially progress more rapidly to a disorder than those who exhibit dependence symptoms initially.

In addition, FMM results suggest that 'problems with the law' is not a reliable indicator for assessing OUDs, which may be related to the reason that its occurrence depends on environmental factors (getting arrested). This criterion is infrequently endorsed (2.5%) by opioid users, and its low discrimination but high severity estimates indicate high measurement errors. This finding is consistent with some results from IRT analyses of DSM-IV criteria for drug use disorders (Langenbucher *et al.* 2004; Lynskey & Agrawal, 2007). Furthermore, DSM-IV uses the presence of either 'tolerance' or 'withdrawal' to distinguish the *physical-dependence* subtype from a *non-physical-dependence* subtype, and the former is considered a more severe group (APA, 2000). Although this subtyping is applied to all substance classes, it has

had little psychometric research. Our mixture analyses that provided both latent categories and dimensional estimates, however, did not find support that 'tolerance' or 'withdrawal' assessed a more severe end of opioid use problems than others, suggesting that this subtyping for dependence may need to be evaluated (Wu *et al.* 2009a).

External validators

Finally, we found preliminary support for the distinction between the more severe and mild groups of non-prescribed users: less educated, greater non-prescribed opioid use, past-year major depression, and greater use of substance abuse treatment were associated with the more severe group as compared with the less severe group.

Limitations and strengths

These findings should be interpreted with some caution. Like other national surveys, NSDUH relies on respondents' self-reports. Although our analysis of past-year OUD criteria and associated behaviors might minimize recall errors, self-reports could be influenced by memory errors and under-reporting. Additional research is needed to evaluate the clinical utility of empirically defined diagnostic groups. In addition, these findings do not apply to a small group (<2%) of incarcerated, institutionalized or homeless adults who are not included in the NSDUH.

NSDUH also has noteworthy strengths. It uses computer-assisted interviewing methods to increase respondents' reporting of sensitive drug use behaviors (Turner *et al.* 1998). The survey also uses a 'pill card' displaying the names and color photographs of opioids to aid respondents' identification and recall of drugs used. Furthermore, its sample is representative of non-institutionalized Americans. It includes the largest sample of non-prescribed opioid users currently available for this research. The results therefore have a high level of generalizability for subgroups.

Conclusions

A factor mixture model combining categorical and dimensional features of classification empirically fits better to diagnostic criteria for prescription OUDs in adults than a pure categorical or dimensional approach. As observed in recent research on prescription OUD criteria in adolescents (Wu *et al.* 2009c), the DSM-IV's hierarchical specification for dependence and abuse is not fully justified for adult non-prescribed opioid users. Research is needed to examine further the utility of this mixture classification for

specific substance use diagnoses and to determine prospectively whether empirically defined diagnostic groups predict treatment response and clinical courses.

Acknowledgments

This work was supported by research grants from the U.S. National Institute on Drug Abuse of the National Institutes of Health: R01DA019623, R21DA027503, R01DA019901 (PI: L.-T. Wu); K05DA017009, U10DA013043 (PI: G. E. Woody); and HSN271200522071C (PI: D. G. Blazer). The sponsoring agency had no further role in the study design and analysis, the writing of the report, or the decision to submit the paper for publication. The Substance Abuse and Mental Health Data Archive provided the public use data files for NSDUH, which was sponsored by the Office of Applied Studies of the Substance Abuse and Mental Health Services Administration. We thank A. McMillan for her editorial assistance and the three anonymous reviewers for their very valuable time and helpful comments.

Declaration of Interest

Dr Woody is a member of the RADARS post-marketing study external advisory group, who assess abuse of prescription medications. Denver Health and Hospital Authority is a non-profit public hospital that administers the RADARS® System, and its costs are supported by contracts with pharmaceutical companies.

References

- APA (2000). *Diagnostic and Statistical Manual of Mental Disorders*, 4th edn, text revision. American Psychiatric Association: Washington, DC.
- Beseler C, Jacobson KC, Kremen WS, Lyons MJ, Glatt SJ, Faraone SV, Gillespie NA, Tsuang MT (2006). Is there heterogeneity among syndromes of substance use disorder for illicit drugs? *Addictive Behaviors* **31**, 929–947.
- Browne MW, Cudeck R (1993). Alternative ways of assessing model fit. In *Testing Structural Equation Models* (ed. K. A. Bollen and J. S. Long), pp. 136–162. Sage Publications: Newbury Park, CA.
- Bucholz KK, Heath AC, Reich T, Hesselbrock VM, Kramer JR, Nurnberger Jr. JI, Schuckit MA (1996). Can we subtype alcoholism? A latent class analysis of data from relatives of alcoholics in a multicenter family study of alcoholism. *Alcoholism: Clinical and Experimental Research* **20**, 1462–1471.
- Feingold A, Rounsaville B (1995). Construct validity of the dependence syndrome as measured by DSM-IV for different psychoactive substances. *Addiction* **90**, 1661–1669.
- Gillespie NA, Neale MC, Prescott CA, Aggen SH, Kendler KS (2007). Factor and item-response analysis DSM-IV criteria for abuse of and dependence on cannabis, cocaine, hallucinogens, sedatives, stimulants and opioids. *Addiction* **102**, 920–930.
- Helzer JE, Bucholz KK, Gossop M (2007). A dimensional option for the diagnosis of substance dependence in DSM-V. *International Journal of Methods in Psychiatric Research* **16** (Suppl. 1), S24–S33.
- Hu L, Bentler PM (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal* **6**, 1–55.
- Kosten TR, Rounsaville BJ, Babor TF, Spitzer RL, Williams JB (1987). Substance-use disorders in DSM-III-R. Evidence for the dependence syndrome across different psychoactive substances. *British Journal of Psychiatry* **151**, 834–843.
- Kuo PH, Aggen SH, Prescott CA, Kendler KS, Neale MC (2008). Using a factor mixture modeling approach in alcohol dependence in a general population sample. *Drug and Alcohol Dependence* **98**, 105–114.
- Kupfer DJ, Regier DA, Kuhl EA (2008). On the road to DSM-V and ICD-11. *European Archives of Psychiatry and Clinical Neuroscience* **258** (Suppl. 5), 2–6.
- Langenbucher JW, Labouvie E, Martin CS, Sanjuan PM, Bavy L, Kirisci L, Chung T (2004). An application of item response theory analysis to alcohol, cannabis, and cocaine criteria in DSM-IV. *Journal of Abnormal Psychology* **113**, 72–80.
- Lo Y, Mendell NR, Rubin DB (2001). Testing the number of components in a normal mixture. *Biometrika* **88**, 778.
- Lynskey MT, Agrawal A (2007). Psychometric properties of DSM assessments of illicit drug abuse and dependence: results from the National Epidemiologic Survey on Alcohol and Related Conditions (NESARC). *Psychological Medicine* **37**, 1345–1355.
- Manchikanti L (2007). National drug control policy and prescription drug abuse: facts and fallacies. *Pain Physician* **10**, 399–424.
- Morgenstern J, Langenbucher J, Labouvie EW (1994). The generalizability of the dependence syndrome across substances: an examination of some properties of the proposed DSM-IV dependence criteria. *Addiction* **89**, 1105–1113.
- Muthén B (2006). Should substance use disorders be considered as categorical or dimensional? *Addiction* **101** (Suppl. 1), 6–16.
- Muthén B, Asparouhov T (2006). Item response mixture modeling: application to tobacco dependence criteria. *Addictive Behaviors* **31**, 1050–1066.
- Muthén BO, Muthén LK (2007). *Mplus: Statistical Analysis with Latent Variables (Version 5.2)*. Muthén and Muthén Inc.: Los Angeles, CA.
- Nelson CB, Rehm J, Ustün TB, Grant B, Chatterji S (1999). Factor structures for DSM-IV substance disorder criteria endorsed by alcohol, cannabis, cocaine and opioid users: results from the WHO reliability and validity study. *Addiction* **94**, 843–855.
- Nylund KL, Asparouhov T, Muthén B (2007). Deciding on the number of classes in latent class analysis and growth

- mixture modeling. A Monte Carlo simulation study. *Structural Equation Modeling: A Multidisciplinary Journal* **14**, 535–569.
- Proudfoot H, Baillie AJ, Teesson M** (2006). The structure of alcohol dependence in the community. *Drug and Alcohol Dependence* **81**, 21–26.
- SAMHSA** (2008). *Results from the 2007 National Survey on Drug Use and Health: National Findings*. Substance Abuse and Mental Health Services Administration, Office of Applied Studies: Rockville, MD.
- SAMHSA** (2009). *Results from the 2008 National Survey on Drug Use and Health: National Findings*. Substance Abuse and Mental Health Services Administration, Office of Applied Studies: Rockville, MD.
- Schatzberg AF** (2010). Why is DSM-5 being delayed? *Psychiatric News* **45**, 3.
- SUDAAN** (2006). *SUDAAN User's Manual, Release 9.0*. Research Triangle Institute: Research Triangle Park, NC.
- Teesson M, Lynskey M, Manor B, Baillie A** (2002). The structure of DSM-IV cannabis use disorders in the community. *Drug and Alcohol Dependence* **68**, 255–262.
- Turner CF, Ku L, Rogers SM, Lindberg LD, Pleck JH, Sonenstein FL** (1998). Adolescent sexual behavior, drug use, and violence: increased reporting with computer survey technology. *Science* **280**, 867–873.
- Wu LT, Blazer DG, Patkar AA, Stitzer ML, Wakim PG, Brooner RK** (2009a). Heterogeneity of stimulant dependence: a National Drug Abuse Treatment Clinical Trials Network study. *American Journal on Addictions* **18**, 206–218.
- Wu LT, Pan JJ, Blazer DG, Tai B, Brooner RK, Stitzer ML, Patkar AA, Blaine JD** (2009b). The construct and measurement equivalence of cocaine and opioid dependences: a National Drug Abuse Treatment Clinical Trials Network (CTN) study. *Drug and Alcohol Dependence* **103**, 114–123.
- Wu LT, Pilowsky DJ, Patkar AA** (2008a). Non-prescribed use of pain relievers among adolescents in the United States. *Drug and Alcohol Dependence* **94**, 1–11.
- Wu LT, Ringwalt CL, Mannelli P, Patkar AA** (2008b). Prescription pain reliever abuse and dependence among adolescents: a nationally representative study. *Journal of the American Academy of Child and Adolescent Psychiatry* **47**, 1020–1029.
- Wu LT, Ringwalt CL, Yang C, Reeve BB, Pan JJ, Blazer DG** (2009c). Construct and differential item functioning in the assessment of prescription opioid use disorders among American adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry* **48**, 563–572.
- Zacny J, Bigelow G, Compton P, Foley K, Iguchi M, Sannerud C** (2003). College on Problems of Drug Dependence taskforce on prescription opioid non-medical use and abuse: position statement. *Drug and Alcohol Dependence* **69**, 215–232.