DSM criteria for tobacco use disorder and tobacco withdrawal: a critique and proposed revisions for DSM-5*

Timothy B. Baker¹, Naomi Breslau², Lirio Covey³ & Saul Shiffman⁴
Department of Medicine, Center for Tobacco Research and Intervention, University of Wisconsin School of Medicine and Public Health, Madison, WI, USA;¹ Department of Epidemiology, Michigan State University, East Lansing, MI, USA;² Department of Clinical Psychology in Psychiatry, Columbia University Medical Center, New York State Psychiatric Institute, New York, NY, USA;³ and Smoking Research Group, University of Pittsburgh, Pittsburgh, PA, USA⁴

ABSTRACT
Aims This paper aims to identify appropriate criteria for tobacco dependence assessment, evaluate relevant research and suggest revisions that may be incorporated into DSM-5. Methods Desirable conceptual and psychometric features of tobacco dependence assessments were identified, including the types of outcomes against which such assessment should be validated. DSM-IV criteria were matched against these criteria and compared with other dependence measures. Results DSM-IV criteria were found to be ambiguous, little used in tobacco research, and have relatively low predictive validity. Other dependence measures were found to have greater validity in the prediction of important dependence features such as relapse likelihood. Strength of urges to smoke on typical smoking days and during abstinence, markers of nicotine intake or frequency of smoking and latency to smoke soon after waking were found to be useful dependence measures. Conclusion The use and utility of DSM-5 will be enhanced by eliminating most DSM-IV criteria and adding new ones based on smoking pattern, smoking heaviness, and the severity of craving during periods of smoking and withdrawal.

Keywords DSM-IV, DSM-5, nicotine dependence, tobacco dependence.

INTRODUCTION
Nicotine dependence is recognized as a medical condition in the Diagnostic and Statistical Manual (DSM) of the American Psychiatric Association [1–3], and Tobacco Dependence is recognized in the International Classification of Diseases [4,5] [NB—for this paper we use the term ‘tobacco dependence’ (or TD) to refer to both].¹ Both systems use checklist criteria to classify individuals as dependent or not dependent (see Table 1 for DSM-IV items) and both claim to be based on empirical data and observations.

This paper reviews and evaluates the criteria comprised by the current DSM-IV and ICD-10 nosologies (for brevity we will use the term DSM to refer to both DSM and ICD criteria, as they are very similar). New criteria are then nominated for DSM-5, along with rationales to support them. Further, it is recommended that most extant DSM-IV criteria be dropped, or at least fundamentally rewritten. The suggested changes are timely because of the current effort to develop new criteria for DSM-5 [6]. In this paper the terms ‘measure’ or ‘assessment’ refer to type of assessment instrument such as the DSM or the Fagerström Test for Nicotine Dependence (FTND). The term ‘criteria’ refers to the ‘assessment domains’ or constructs targeted by a measure (e.g. tolerance), as distinct from the measure itself. ‘Items’, on the other hand, refers to the specific questions designed to assess a

¹We will use the term tobacco dependence versus nicotine dependence because in the vast amount of dependent use, tobacco is the vehicle for nicotine delivery, and there is no clinical utility in speculating whether a particular tobacco user’s dependence is attributable to nicotine per se. Note that the DSM-5 may use the term ‘substance use disorder’ to refer to diagnosable drug use disorders [6].

*The authors have no formal role in the development of DSM-5 criteria; this paper expresses their own views and not those of any official body.
criterion or construct. Finally, the term ‘outcomes’ refers to the behaviors or variables used to test the validity or accuracy of a measure or item (what it is intended to predict).

**Rationale for a Revision of the DSM Criteria**

The DSM criteria, presented in Table 1, need significant revision. Evidence shows that the criteria and the scoring strategies recommended in recent versions of the DSM have not performed well relative to other measures of tobacco dependence. (A similar charge can be laid against the ICD-10 criteria, because the two systems are highly similar [7]). Researchers have produced alternative instruments to assess dependence, such as the FTND, the Nicotine Dependence Symptom Scale (NDSS), the Wisconsin Inventory of Smoking Dependence Motives (WISDM) and the Hooked on Nicotine Checklist (HONC) [8,9]. These, especially the FTND, have achieved much wider use than the DSM criteria, probably because of their ease of use and their superior prediction of important outcomes.

The goals and nature of the DSM diagnosis of tobacco dependence

Diagnosis of tobacco dependence should help to inform treatment decisions or recommendations, identify individuals who are at risk for negative consequences from tobacco use, identify those who need intervention in order to quit tobacco use successfully and suggest what type or ‘dose’ of treatment to use (see [10]). Finally, a dependence measure should be useful in research.
DESIRABLE FEATURES OF A DEPENDENCE ASSESSMENT

Growing out of these goals of diagnosis, the authors believe that an ideal TD assessment should have the following features (not listed in order of importance):

• The criteria should yield items that are very clear and subject to straightforward scoring. The information sought by each question should be clear to both the assessor and to the respondent, minimizing the need for extensive interpretation. Moreover, response options should be relatively few in number and converted easily to a suitable score.

• The criteria should address constructs or domains that are theoretically and empirically central to dependence.

• The assessment should be brief, making it appropriate for use in clinical settings.

• The criteria should yield items and measures with sound psychometric properties; specifically, they should be reliable and valid. Reliability refers to the repeatability of results (e.g. obtaining the same diagnosis or score) when the underlying TD is stable. Validity relates to prediction of important outcomes seen as centrally relevant to dependence, such as future likelihood of quitting smoking successfully; heaviness of present and future use; need for, or response to, dependence-focused smoking cessation treatment. Other outcomes could certainly be used, such as the tendency to experience withdrawal symptoms upon drug discontinuation, or strong relations with molecular genetic substrata of heavy tobacco use.

The American Psychiatric Association (APA) DSM-IV attempts to achieve the above goals via assessment of seven dependence criteria perceived to co-occur with problematic drug use; see Table 1: e.g. presence of tolerance to tobacco effects, occurrence of withdrawal symptoms, difficulty controlling or stopping tobacco use and so on. Diagnosis requires the clinician to determine whether three or more of the seven main criteria (Part 1) occur simultaneously within a 12-month period. There is no clear guidance as to how the clinician acquires this information. Therefore, diverse item wordings and delivery methods have been used in DSM-based assessment.2 This creates some uncertainty about whether research results reflect the worth of the items, the criteria themselves or both.

EVALUATING DSM-IV WITH REGARD TO CORE OUTCOMES

Below we evaluate the DSM-IV criteria against the ideal features listed above.

Table 2 Fagerström Test for Nicotine Dependence (FTND) items.

<table>
<thead>
<tr>
<th>Item</th>
<th>Response options</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. How soon after waking do you smoke your first cigarette?</td>
<td>• within 5 minutes/6–30 minutes</td>
</tr>
<tr>
<td></td>
<td>• 31–60 minutes/after 60 minutes</td>
</tr>
<tr>
<td>2. Do you find it difficult to refrain from smoking in places where it is forbidden?</td>
<td>• Yes/no</td>
</tr>
<tr>
<td>3. Which cigarette would you hate to give up?</td>
<td>• The first one in the morning</td>
</tr>
<tr>
<td></td>
<td>• All the others</td>
</tr>
<tr>
<td>4. How many cigarettes do you smoke?</td>
<td>• 10 or less/11–20/21–30/31 or more</td>
</tr>
<tr>
<td>5. Do you smoke more frequently during the first hours after waking than during the rest of the day?</td>
<td>• Yes/no</td>
</tr>
<tr>
<td>6. Do you smoke if you are so ill you are in bed most of the day?</td>
<td>• Yes/no</td>
</tr>
</tbody>
</table>

Note. See [27].

• The criteria should yield items that are very clear and subject to straightforward scoring.

We believe that the DSM-IV criteria do not provide for clear items or scoring. For instance, criterion 4 asks about ‘persistent desire to cut down...or control tobacco use’—but the definition of ‘persistent’ could vary greatly from one clinician or one patient to another. Criterion 6 addresses ‘important social, occupational, or recreational activities...given up or reduced because of nicotine’. Applying this to smoking is difficult, as one can often smoke during social, occupational and recreational activities. Conversely, if one interprets the criterion broadly, to include, e.g. missing part of a movie or a meeting in order to smoke, it is hard to imagine any regular smoker for whom this has not been true. Thus, interpretation is crucial for this criterion: what makes the activity important and how much must participation be reduced? The ambiguity of the DSM criteria can be compared with some of the items of the FTND, an alternative dependence instrument; the FTND items, by comparison, seem much more straightforward (e.g. items 1 and 4; see Table 2).

• The criteria should address constructs that are central theoretically and empirically to dependence.

Recent research has identified several characteristics of tobacco use patterns that appear to be core dependence features, including tobacco use that:

2Although the DSM does provide some guidance as to how the criteria may apply to nicotine use. Further, the points made in this paper should also apply to the diagnosis of tobacco use disorder which is a less severe form of tobacco use disorder than nicotine dependence that is proposed for DSM-5: see www.dsm5.org (Archived at http://www.webcitation.org/62nzhVtw6 on 29 October 2011.)
withdrawal is thought to be a core feature of dependence, and one that is relatively independent of heaviness of smoking. DSM-IV does ask about experience of prior withdrawal. However, the validity of this criterion may be limited as it does not assess withdrawal-related craving (see below).

Three of the DSM-IV criteria (5, 6 and 7) focus on the ‘costs’ of smoking or other drug use. Continuing the behavior despite costs may reflect how compulsive the behavior is. However, the costs assessed by the DSM criteria will vary significantly according to the person’s life context and the extent to which they might normally limit tobacco use; e.g. diagnosis is more likely for those who cannot afford tobacco, who are engaged in employment that prohibits smoking or who have a tobacco-related disease. While continuing to smoke in the face of such costs makes dependence more evident, dependence surely exists even without such costs, or awareness of such costs. Thus, focusing on costs makes the diagnosis too highly bound to an individual’s social and health context. Was tobacco dependence non-existent in the 1940s when tobacco use was thought to be innocuous, when it little disturbed the fabric of life and when quit attempts were rare (why quit)? Are low-income smokers more dependent because the costs of purchasing tobacco weigh more heavily on them? Would smoking become any less addictive if it were somehow rendered harmless?

An alternative view is that dependence is produced by biological and psychological processes that do not depend so thoroughly on social context and, indeed, may be manifested by a rat in a cage: e.g. high rates of self-administration, withdrawal upon falling blood levels of nicotine, heightened resistance to extinction, withdrawal-induced anhedonia, high probability of relapse, rapid reinstatement of drug use after abstinence, strong conditioned reinforcement of tobacco use, relative insensitivity to the aversive consequences of self-administration (‘costs’, again, in a clearer conceptual context) and so on. We argue that criteria that assess such core features more directly are likely to result in more valid and useful measures.

- The assessment should be brief, making it appropriate for use in clinical settings.

The proposed DSM-5 will, apparently, include 11 criteria (http://www.dsm5.org; archived at http://www.webcitation.org/62nzhVtw6 on 29 October 2011), which may be unnecessarily long—both in terms of criteria and time required. Two current measures—the FTND (five items) and the HSI [23] (two items from the FTND)—are shorter and have good reliability and validity. Importantly, both are amenable to rapid self-administration, so take considerably less time than a DSM-IV assessment. Thus, for clinicians to use longer instruments such as ones based on the DSM, the instruments will have the burden of showing incremental validity or utility over briefer ones, and most data suggest this is not the case (e.g. [7.28–30]).

- A tobacco dependence diagnostic measure should be both reliable and valid, allowing meaningful prediction of: future likelihood or ability to quit smoking successfully; heaviness of present and future use; response to smoking cessation treatment; and harm or costs incurred by tobacco use, including occupational, social, health and psychological or psychiatric harms.
Reliability

The DSM-IV criteria appear to have acceptable reliability. There are two types of reliability. The first focuses on internal consistency across the constituent items, essentially reliability across items on a single administration. There is evidence that the DSM criteria generate items that are related to a single construct [31–33], although a unifactorial solution may not be optimal [11,34]. There is evidence that the interviewer administered format is less internally consistent ($\alpha = 0.50$ with the DIS-interviewer administered format) [35] than paper-and-pencil formats ($\alpha = 0.80$) [36]. The second type of reliability is test–retest reliability, and concerns the repeatability of the test performance over time. Importantly, there is evidence that DSM-IV-based assessments have acceptable test–retest reliability and that the paper-and-pencil and structured interview administration formats are meaningfully related (0.69–0.79) [37]. A caveat is that stability across repeated testing (test–retest reliability) is not always a desirable characteristic in a measure, because stability can be achieved by making the measure insensitive to change (consider the case of using age-of-onset of smoking as a measure of dependence). An ideal measure is stable when the underlying construct has not changed, but changes when there are genuine changes in the underlying condition. Little is known about the ability of DSM or other measures to reflect genuine shifts in dependence.

Validity

Valid measures predict key outcomes related to the construct that they are intended to assess. DSM-based measures do not impressively predict major dependence outcomes. Such measures yield some evidence of predictive validity: scores on the Tobacco Dependence Screener (the TDS, a measure based on DSM [36]) are related to measures of tobacco self-administration (e.g. number of cigarettes smoked/day or light versus heavy smoking), can distinguish between light and heavy smokers [36,38] and can predict increased subsequent smoking among neophyte adolescent smokers [39,40]. Also, at least one study reports that DSM-type items predict cessation likelihood as well as or better than other dependence measures such as the FTND [36].

However, most research is less favorable, showing that alternative TD measures have superior predictive validity, certainly for outcomes such as smoking heaviness and cessation likelihood (e.g. [26,29,35]). We believe that a measure of TD should be expected to predict three key features or outcomes of TD: withdrawal severity, heaviness of use and cessation of tobacco use. Withdrawal is an important outcome as it is a characteristic feature of all dependence disorders (a sine qua non), it is an intrinsic cost of addiction and it plays a significant role in preventing reduction or discontinuation of tobacco use [41,42]. Smoking heaviness is important, as a high level of involvement with a drug is linked conceptually and empirically with dependent use [19] and it is highly related to probability of harms due to drug use (e.g. health effects). Finally, it seems especially vital that a TD measure predicts likelihood of future cessation, as a key facet of the dependent drug use is its intractability. Accordingly, we assess the validity of DSM and other measures by their association with these three constructs, which are moderately associated with one another (e.g. [43,44], although cf. [45]), suggesting that they reflect related, but different, aspects of the dependence construct.

Research with both clinical and population-based samples yields a fairly consistent pattern in which DSM-IV-based measures are less predictive of smoking cessation than are other types of dependence measures [28,29,35,46]. For instance, the DSM criteria (assessed with the NIMH-DIS) were less predictive of smoking status over a 3-year longitudinal study than was the FTND [28]; also cf. [47]. The items of the FTND were superior whether both measures were used as binary (dependent/non-dependent) or continuous measures. In addition, multiple clinical trials have shown that the FTND predicts cessation success better than does the DSM [9,29,35,46] [with odds ratio (ORs) = 2–3 [22]]. In addition, existing research suggests the DSM-IV-based instruments tend to have modest relations with both smoking heaviness measures and withdrawal [29,47], relations seemingly weaker than those of alternative dependence instruments [12,29].

SUGGESTED CRITERIA FOR THE ASSESSMENT OF TOBACCO DEPENDENCE

Proposed criteria for DSM-5 and associated sample items are presented below along with rationales for their inclusion (see Table 3). Criteria are recommended based upon: (i) consistency with theory or conceptual perspectives on TD; (ii) evidence that the criterion indexes heavy tobacco use; (iii) evidence that the criterion indexes likelihood of cessation; and (iv) associations with other measures that are linked conceptually or empirically with dependence. Further, the current recommendations are based on the view that the many differences between tobacco use and other types of dependent drug use (see above) support a diagnostic strategy that is tailored to tobacco per se, rather than one designed to be appropriate for all drugs [48]. Finally, as suggested above, any DSM-5 criteria should foster measures that are relatively brief, the items should be clear and easy to answer and appropriate for use in diverse clinical contexts.
Table 3 Recommended criteria and example items for DSM-5.

(1) Criterion: strong and bothersome craving both during on-going smoking and when quitting or reducing smoking
Sample items:
‘When I’m really craving a cigarette, it feels like I’m in the grip of some unknown force that I cannot control’ (NDSS)
‘I frequently crave cigarettes’ (Brief WISDM)

(2) Criterion: Significant withdrawal symptoms, including craving, during smoking cessation or reduction.
‘Whenever I go without a smoke for a few hours, I experience . . . ’ (NDSS)
‘When I haven’t been able to smoke for a few hours, I notice . . . ’ (Brief WISDM)
(a) Craving
(b) Irritability, anger
(c) Anxiety
(d) Sadness, depression
(e) Difficulty concentrating
(f) Insomnia
(g) Impatience
(h) Restlessness

(3) Criterion: latency to smoke upon awakening.
Sample item: ‘On days that you can smoke freely, how soon after waking do you smoke your first cigarette?’ (FTND)
(i) Within 5 minutes
(ii) 6–30 minutes
(iii) 31–60 minutes
(iv) After 60 minutes

(4) Criterion: smoking heaviness; average number of cigarettes smoked/day on a typical day of smoking
Sample item: ‘On days that you smoke, how many cigarettes do you usually smoke?’ (FTND)
(i) 10 or less
(ii) 11–15
(iii) 16–19
(iv) 20–25
(v) 26–30
(vi) 31 or more

The sample items are derived and adapted from items contained in either the Fagerstrom Test for Nicotine Dependence (FTND), the Nicotine Dependence Symptom Scale (NDSS) or the Wisconsin Inventory of Smoking Dependence Motives (WISDM and Brief WISDM).

Two other issues are considered: (i) breadth of the TD construct, because too narrow a focus may boost some indices (e.g. internal consistency) but may reduce overall utility [49,50] and (ii) ‘difficulty’ of an item or criteria. An ‘easy’ item or scale (one that is ‘passed’ by many individuals) may be sensitive to low levels of the construct, but may lack both specificity and discrimination at higher levels of dependence. An item that is too ‘hard’ may set the bar too high and miss detection of meaningful levels of dependence (i.e. good specificity but low sensitivity). The ‘difficulty level’ of the DSM, and its possible implications for sensitivity and specificity, is a vital question. The field has not achieved consensus on the proper ‘difficulty level’ for optimal TD diagnosis [51]. Because virtually all regular or daily smokers find it difficult to quit, experience withdrawal and face huge personal costs from smoking, one could argue that almost all smokers should be diagnosed as dependent (see [10] for related concerns); however, only about half of daily smokers meet DSM diagnoses for dependence. Conversely, it is clear that among daily smokers there is considerable variability in withdrawal severity, relapse risk and tobacco exposure [22]. One potential option is to adopt a quasi-continuous or continuous index of dependence [52].

Table 3 presents the criteria, and sample items that could be used to assess them, in a new DSM-5. Two of the sample items are adapted fairly directly from the FTND, while examples of other items are derived from questionnaires that have successfully predicted dependence outcomes (the NDSS and WISDM and Brief WISDM [53]. These items are intended as examples; obviously, the optimal form of a DSM-5 criterion, and any implementation into items and their response and scoring properties, would have to be determined via further research.

Rationale for included criteria

Criterion 1: craving

Tobacco craving reflects a conscious desire or intention to use tobacco. Many accounts of TD include craving as an important dependence symptom [15,54–56], and there is evidence that craving plays a causal role in relapse [14,41,57–60], is correlated meaningfully with smoking heaviness [17,19,25] and is correlated with other important features of TD [29,61]. Among all withdrawal symptoms, craving stands out as being associated most highly with TD, and as a predictor of abstinence following a quit attempt [14,62–64]. Finally, subscales that assess craving tend to be the most valid subscales on the major multi-dimensional measures of TD; i.e. NDSS ‘drive’ (e.g. [18]) and WISDM ‘craving’ (e.g. [9]).

The craving items reproduced in Table 3 are not specific to withdrawal craving. This is because while deprivation typically increases craving [65], craving also occurs during ongoing smoking [62,66], and predicts smoking satisfaction [67,68] and future ad libitum smoking [66,69]. Withdrawal-related craving is targeted by the recommended withdrawal criterion (see Table 3, criterion 2). Finally, inclusion of a craving criterion harmonizes the DSM and ICD criteria (as the latter includes compulsion-to-use as a criterion).

Criterion 2: withdrawal

Including withdrawal in the dependence criteria is consistent with most theories of dependence (e.g. [57,70–73]; see [74]). There is a strong association between withdrawal and future cessation success ([41,57,59,75–78]; but see [74]). Also, withdrawal severity indexes smoking
heaviness across the ontogeny of TD [40, 62, 79, 80] appears to distinguish heavy, regular tobacco users from light or intermittent users [81, 82], and is related significantly to TD [43, 44, 83]. In addition, withdrawal is aversive and therefore constitutes a problem or cost of dependent tobacco use that the dependent smoker must endure if s/he persists in such use [74].

Craving was not included in the DSM-IV withdrawal symptoms because of evidence that it sometimes does not increase in response to a discontinuation of tobacco use [84]. This may reflect the fact that the rise-time of craving after smoking discontinuation is often so rapid that fine temporal resolution is needed to detect it [65, 85]. However, most research shows that craving reports typically escalate in response to tobacco abstinence and are associated significantly with other elements of that syndrome [78, 86–89]. In addition, as noted earlier, craving is the withdrawal symptom that is most highly predictive of cessation success [14, 19, 59, 75, 77]. Thus, DSM-5 should assess craving that occurs as part of the withdrawal syndrome (criterion 2), and more generally as well (criterion 1).

Criteria 3 and 4: time to first cigarette (TTFC) after waking and usual number of cigarettes consumed per day (CPD)

These criteria are based on FTND items (items 1 and 4). There is evidence that the FTND and DSM assessments tap somewhat distinct domains; e.g. the correlation between the two is fairly modest. For instance, in one study the correlation between the FTND and a paper-and-pencil version of the DSM-IV was only 0.26 [90]. Similarly, the FTND was correlated with DSM-IV diagnosis determined via structured interview at only r = 0.24–0.35 [91]. Thus, there is only a modest level of shared variance between the two measures and they diagnose somewhat different groups of people as dependent ([92]; also see [28, 91, 93, 94]).

As noted earlier, numerous researchers have found that the FTND can predict cessation likelihood ([28, 91, 93, 96]; but see [97]). Moreover, there is evidence that smokers with higher FTND scores show greater benefit from nicotine replacement therapy (NRT) than do those with low scores [98]. The FTND tends to show only modest predictive validity with regard to ongoing ratings of withdrawal severity [27]; however, attitudinal factors may account for stronger relations when global impressions of withdrawal are used [96, 99].

The FTND asks directly about smoking heaviness (both items 1 and 4 relate to this: Table 2). There is potential for circularity when an assessment asks about smoking heaviness and then validates it with a self-report of smoking heaviness. This applies especially to item 4, but less so to item 1. It should be noted, however, that these items correlate with both carbon monoxide and cotinine levels [8, 99–101], which are related linearly to blood nicotine level, demonstrating that the items index effective dose or metabolic biomarkers. Secondly, both items predict likelihood of future cessation (e.g. [22]): thus, their predictive validity cannot be attributed to common method variance or misleading self-perception. The evidence suggests that these items assess directly and economically a clinically and theoretically meaningful TD dimension.

Direct comparisons of the FTND- and DSM-based measures show that the former is related more strongly to the prediction of smoking heaviness and relapse vulnerability. For instance, the FTND shows stronger relations with both self-report and biochemical measures of smoking heaviness than do DSM-type measures (e.g. [9, 91]), and shows stronger relations with cessation likelihood (e.g. [22, 28]).

Research by Heatherton and his colleagues [23] found that among FTND items, 1 and 4 (TTFC and CPD) were particularly predictive of effective dose biomarkers (e.g. carbon monoxide, blood/urine nicotine and cotinine levels) and labeled these items the Heaviness of Smoking Index (HSI: [8]). These findings agree with earlier positive findings on similar items [102–105]. Subsequent research has generally supported the conclusion that the two HSI items account for much of the discriminative and predictive validity of the FTND with regard to both effective dose biomarkers and cessation likelihood [22, 106]. The two HSI items are also being used for assignment of dosage of nicotine replacement treatment [107]; see [108]. Finally, the two items are correlated strongly but can yield additive predictive validity (cf. [22, 28, 35, 95]), suggesting that both items be retained to provide meaningful coverage of this important domain.

A recent analysis tested whether adding a cigarette quantity–frequency item, reflecting daily smoking rate, increased model fit when added to the set of DSM-IV criteria [109]. The results suggested that a CPD item actually decreased model fit. This suggests that adding a single item to the existing DSM criteria is probably not a good strategy for improving the DSM (although see [32]): greater changes to the DSM criteria are needed.

The four criteria listed above tend to have a stronger evidence base than any of the criteria in DSM-IV, and they should certainly be tested for inclusion in DSM-5. However, other criteria should also be tested for possible inclusion. For instance, we believe that two criteria from the DSM-IV should be modified to make them more appropriate for tobacco use, and then tested for possible inclusion in the DSM-V, namely:

- Does the person spend a great deal of time in activities in order to obtain or use tobacco (e.g. chain smoking, moving to locations where smoking is permitted,
making efforts to buy cigarettes, obtaining money for cigarettes) or recover from its effects (get over illness)?

- Does the person give up or change social, work or recreational activities because of tobacco use (activities are affected by the need to smoke; the need to smoke interrupts work)?

There is little evidence on the relative validity of individual DSM-IV items. However, the two constructs noted above are conceptually congruent with dependence [26,35], and re-wording to make them more focused on TD (such as was employed above) may enhance their validity. The wording above is just for purposes of illustration.

There is even less evidence supporting the inclusion of the other DSM-IV criteria: desire to, or difficulty in, cutting down or controlling use; continued use despite harm; tolerance or diminished effect with continued use; use for withdrawal relief; and use of tobacco in larger amounts than intended (see Table 1). Such criteria have not been related meaningfully to core outcomes [26,28,35], and some are double-barreled or require difficult judgments (cf. [110]).

Finally, while it might be attractive to add criteria because they seem conceptually appropriate, each added criterion comes with a potential cost—not only in terms of burden, but also because each added criterion has the potential to dilute the proportion of variance in the instrument (assessment) that reflects more valid criteria or items.

**DISCUSSION**

Considerable evidence suggests that the four core criteria displayed in Table 3 should be represented in DSM-5 if it is to possess optimal validity. Therefore, these criteria should definitely be included in any test versions of DSM-5. Any thorough and meaningful development of a TD diagnostic assay should further evaluate the following with regard to criteria and items derived from them (in the case of DSM based questionnaires): (i) cut-scores for multi-category items; (ii) cut-scores, weights or measurement and response option models for scores for the entire set of criteria or items; (iii) item response theory (IRT) analyses that evaluate the relation of the criteria or items with the new TD latent variable; (iv) measurement invariance analyses that address criteria, item and measure performance across critical populations; and (v) need for further criteria and item trimming. It is especially important to determine which criteria or items yield incremental validity when the best performing ones have been entered into prediction models. While some validity data are available on relations of individual DSM criteria with cessation outcomes or smoking heaviness measures [26,35], such relations do not reveal which criteria may contribute to validity in the context of other valid criteria. It is also vital to examine the relation of new criteria with additional TD outcomes: treatment outcomes and contribution to treatment algorithms, sensitivity to critical gene variants, work for nicotine/tobacco in instrumental self-administration paradigms and so on.

The listed candidate criteria (Table 3) do not exhaust the list of potential good candidates for inclusion. For instance, some evidence would support the development of new candidate criteria that directly elicit information about the smoker’s subjective sense of loss of control over tobacco (see [7,19,46]), waking up to smoke [111,112] or simple self-ratings of degree of dependence [113]. Criteria tapping perceived loss of control seem to be related especially highly to TD outcomes [19,114]. A criterion that focuses on longest duration of abstinence after a prior quit attempt might provide a means of asking about prior exercise of control over smoking (e.g. [115,116]).

A vital question regarding the use of any new DSM assessment of TD concerns how the scale score (criteria or item endorsement) is translated into a categorical appraisal of dependence (e.g. 0–1 = not dependent, 2–4 = dependent; >4 = severe dependence). It is currently unclear what sort of dependence metric is optimal for TD (see [10]). Presumably, empirically driven strategies should be used to generate appropriate cut-scores [e.g. decision-tree models, receiver operating characteristic (ROC) curve methods; e.g. [117]]. In addition, a dimensional and continuous metric could be developed for diagnosis, although some research [17] suggests that dependence may be taxonic (possessing features of a discrete category). The criteria and research methods we have outlined would be appropriate for the development of either a dimensional or quasi-continuous scoring system, or for a continuous scoring system for research and a categorical system for clinical use [118].

We believe there is also a need to rethink the functionality of a nicotine dependence diagnosis [10]. Under DSM-III-R diagnostic criteria and cut-offs, only about half of daily smokers earned diagnoses in some age cohorts [119]. However, smoking causes major health risks to all daily smokers, and even daily smokers scoring at low levels on dependence criteria have difficulty quitting [120,121], may benefit from cessation intervention [122,123] and suffer from the numerous costs of smoking (economic, loss of time, etc.). Thus, the current criteria and diagnostic rules may be set at too high a threshold. As opposed to agents such as alcohol, tobacco does not foster widespread, relatively innocuous use. Achieving an optimal measure cut-score for TD should be one component of a greater effort to improve the utility and clinical benefit of diagnosis.

Using a smoking rate criterion in DSM-5 (cigarettes smoked/day) may raise concerns because smoking rates...
may change over time, and may thereby change the relation between the measure and dependence outcomes. While this would introduce complexity, it might not necessarily compromise the utility of such criteria or items. First, as an ebbing tide may lower all boats, relative differences in smoking heaviness might continue to predict dependence well, despite population shifts in heaviness. Of course, this would argue for periodic re-norming of cut-scores, or using items that do not require information on specific numbers of cigarettes consumed (e.g. self-perception of smoking heaviness and the automaticity of smoking [27]).

In conclusion, substantive and psychometric considerations strongly support four new criteria for further testing and potential inclusion in the DSM-5. These criteria concern smoking heaviness, latency to smoke upon awakening, craving to smoke and withdrawal severity. Other criteria were identified for additional research and possible inclusion. Little evidence supported the inclusion of most existing DSM-IV items, but re-wording and re-formulation might improve their validities. This work also nominated important outcomes for the evaluation of TD criteria or items, and suggested important topics or guiding principles to advance the assessment of TD; e.g. regarding the difficulty level of the assessment (the severity of use that warrants diagnosis) and whether the assessment should be specific to tobacco dependence versus applicable to any substance use disorder. The strongest recommendation is that the development of DSM-5 be driven by systematic, psychometric evaluation with regard to theoretically and clinically meaningful criteria. The DSM-IV has face validity because it directly assesses occurrence of harm as a function of tobacco use; this virtue is outweighed by its poor predictive validity. More research is clearly needed on this important topic.

Declarations of interest

Timothy B. Baker is supported by NIH grant P50 DA019706-10; Naomi Breslau is supported by NCI grant P01 CA89392; Lirio Corey is supported by NIDA grant U10 DA01332-83; Saul Shiffman is supported by NIDA grant R01 DA02074.

Acknowledgements

The authors thank Dr Wendy Theobald for insightful comments and assistance. John Hughes provided consultation and feedback on this project.

References

10. West R., Miller P. What is the purpose of diagnosing addiction or dependence and what does this mean for establishing diagnostic criteria? Addiction 2011; 106: 863–5.


