

Screening for cannabis use disorder among young adults: Sensitivity, specificity, and item-level performance of the Cannabis Use Disorders Identification Test – Revised

Sophie G. Coelho¹, Christian S. Hendershot^{2,3}, Lena C. Quilty^{4,5}, Jeffrey D. Wardell^{1,5,6*}

¹Department of Psychology, York University

²Bowles Center for Alcohol Studies, University of North Carolina–Chapel Hill, Chapel Hill, NC

³Department of Psychiatry, University of North Carolina–Chapel Hill, Chapel Hill, NC

⁴Campbell Family Mental Health Research Institute, Centre for Addiction and Mental Health, Toronto, Canada

⁵Department of Psychiatry, University of Toronto

⁶Institute for Mental Health Policy Research, Centre for Addiction and Mental Health, Toronto, Canada

*Corresponding author: Jeffrey D. Wardell, 277 Behavioural Sciences Building, York University, 4700 Keele St., Toronto, ON, Canada, M3J 1P0, Phone: 416-736-2100 x. 44241, Email: jwardell@yorku.ca

Declarations of Interest: None.

Funding: This research was supported by the Caskey/Francis Family Award in Clinical Research awarded by the Centre for Addiction and Mental Health Foundation (PIs: Jeffrey D. Wardell & Christian S. Hendershot). Funding sources had no role in study design; the collection, analysis, and interpretation of data; the writing of the report; and in decision to submit the article for publication

Acknowledgements: The authors wish to thank Dinat Khan and Brian Pereira for their assistance with data collection for the study.

Formal publication: Coelho, S. G., Hendershot, C. S., Quilty, L. C., & Wardell, J. D. (2024). Screening for cannabis use disorder among young adults: Sensitivity, specificity, and item-level performance of the Cannabis Use Disorders Identification Test - Revised. *Addictive behaviors*, 148, 107859. <https://doi.org/10.1016/j.addbeh.2023.107859>

© 2024. This manuscript version is made available under the CC-BY-NC-ND 4.0 license <https://creativecommons.org/licenses/by-nc-nd/4.0/>

Highlights

- The CUDIT-R was accurate in identifying young adults with cannabis use disorder.
- A cut-point of nine yielded optimal sensitivity and specificity.
- Items varied in difficulty and most items displayed moderate to high discrimination.
- Items assessing hours high and physically hazardous use displayed poor discrimination.
- The CUDIT-R may be useful in screening for cannabis use disorder among young adults.

Abstract

The Cannabis Use Disorders Identification Test Revised (CUDIT-R) is a widely used screening tool. However, its utility in screening for cannabis use disorder (CUD) among young adults requires further investigation. The current study evaluated the accuracy of the CUDIT-R in distinguishing between young adults with and without CUD. We also examined the individual item properties of the CUDIT-R in this sample. A community sample of young adults ($N=153$, ages 19–26 years) who reported at least one lifetime period of regular cannabis use (i.e., twice or more per month for at least six months) completed the CUDIT-R and the Structured Clinical Interview for DSM-5 (SCID-5) Substance Use Module. The CUDIT-R showed good accuracy in distinguishing between young adults with and without CUD as determined by the SCID-5 (area under the receiver operating characteristic curve = 0.84). A cut-point of nine yielded 87.80% sensitivity and 70.42% specificity, with positive and negative predictive values of 77.42% and 83.33%, respectively. Item response theory analyses revealed that CUDIT-R items assessing impaired control, failure to meet expectations, time devoted to use, and memory and concentration problems displayed high discrimination and were relatively difficult, whereas items assessing frequency of use and attempts to reduce use showed moderate discrimination and relatively low difficulty. Items assessing hours high and use with potential for physical hazard displayed poor discrimination. The CUDIT-R appears to validly distinguish between young adults with and without CUD and may be useful in screening for CUD in this high-risk population. Further research considering individual item properties is needed to refine the CUDIT-R for use among young adults.

Keywords: Marijuana; cannabis use disorder; SBIRT; sensitivity; specificity; item response theory; emerging adults

1. Introduction

Cannabis is widely used among young adults in North America, with approximately 35–49% reporting past-year use (Government of Canada, 2021; Substance Abuse and Mental Health Services, 2021). Regular cannabis use confers risk of cannabis use disorder (CUD), which is especially prevalent among young adults (Han et al., 2019; Leung et al., 2020). Yet, young adults are less likely to seek treatment for CUD relative to older age groups (Wu et al., 2017). A Screening, Brief Intervention, and Referral to Treatment (SBIRT) approach may therefore be warranted to help connect young adults with CUD to services (Gette et al., 2023). Such an approach requires screening tools that are effective in identifying at-risk young adults.

The Cannabis Use Disorders Identification Test – Revised (CUDIT-R; Adamson et al., 2010) is the most widely used instrument for assessing hazardous cannabis use. The CUDIT-R includes eight self-report items that assess cannabis consumption, cannabis-related behaviours and consequences, and cannabis dependence symptoms. These items are summed to provide a continuous index of problematic or hazardous cannabis use, with scores ranging from 0 to 32. The original validation study found that a CUDIT-R cut-point of 13 demonstrated high sensitivity (91%) and specificity (90%) in detecting cannabis dependence, as determined by the Structured Clinical Interview (SCID-4) for the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV). However, cannabis dependence is no longer recognized by the DSM-5 (American Psychiatric Association, 2013), which instead characterizes CUD along a severity continuum based on the number of symptoms endorsed (American Psychiatric Association, 2013). The ability of the CUDIT-R to distinguish between those with and without CUD, as characterized by current diagnostic criteria, thus warrants further evaluation. Moreover, the original validation study evaluated the CUDIT-R in a sample of treatment-seeking adults.

Thus, there remains a need to examine the utility of the CUDIT-R as a CUD screening tool for young adults in the community, as they represent a high priority target population for SBIRT efforts.

Although the factor structure of the CUDIT-R among young adults has been established (Risi et al., 2020), only one previous study to our knowledge examined the validity of the CUDIT-R in identifying CUD according to DSM-5 criteria among young adults (Schultz et al., 2019). In this study, the CUDIT-R showed excellent area under the receiver operating characteristic curve (0.91) and a cut-point of 6 was sensitive and specific in detecting CUD among young adult college students reporting past-30-day cannabis use (Schultz et al., 2019). However, this study used a self-report checklist of DSM-5 CUD symptoms administered online, rather than a clinical interview, to determine the presence or absence of CUD. Consequently, the ability of the CUDIT-R to distinguish between young adults with and without DSM-5 CUD diagnoses, as determined using gold standard diagnostic assessments, remains unknown.

The validity of the CUDIT-R in detecting DSM-5 CUD among young adults has also yet to be studied in a jurisdiction in which non-medical cannabis is fully legal. This is a limitation given differences in population-level cannabis use patterns across legal contexts; that is, cannabis is more commonly used and is used at higher frequencies under more permissive cannabis legislation (Goodman et al., 2020; Hammond et al., 2021). In fully legal contexts, higher rates of regular cannabis use among young adults may engender elevated rates of CUD yet pose unique challenges to detecting CUD given that cannabis use is normalized, and thus regular use may be endorsed by both problem and non-problem users. Moreover, items assessing hazardous cannabis use may perform differently in fully legal contexts, in which norms surrounding young adult cannabis consumption and perceptions of risk for harms may differ (Goodman &

Hammond, 2022). A re-evaluation of the CUDIT-R's validity in detecting CUD and individual item properties among young adults in jurisdictions in which cannabis is fully legal is thus needed.

The current study evaluated the performance of the CUDIT-R in a sample of young adults who reported a history of regular cannabis use. The study was conducted in Canada, following the legalization of non-medical cannabis use. We used receiver operating characteristic analyses to examine the accuracy of the CUDIT-R in distinguishing between young adults with and without CUD diagnoses obtained using a gold standard structured clinical interview. We also sought to determine an optimal CUDIT-R cut-point to detect potential CUD. Based on previous research (Adamson et al., 2010; Schultz et al., 2019), we hypothesized that the CUDIT-R would be accurate in distinguishing between young adults with and without CUD. Further, we conducted an item response theory analysis to explore the psychometric properties of individual items of the CUDIT-R in our sample.

2. Method

2.1. Participants and Procedures

The study took place in Toronto, Canada from March, 2019 to March, 2020, after non-medical cannabis legalization. Participants were young adults (ages 19–25 years) of legal age to purchase cannabis who participated in a larger study examining associations between cannabis use and cognitive functioning. Young adults were recruited via advertisements posted on public message boards across different neighborhoods in Toronto, on and around college and university campuses in Toronto, and online to reach as broad of an audience as possible. Interested individuals were invited to complete an online eligibility screener. For the larger study, eligible participants were young adults aged 19–25 who reported at least one lifetime period of regular

cannabis use (defined as using at least two times per month for six months). Exclusion criteria were: (i) history of severe head injury; (ii) current or past treatment for alcohol or substance use; (iii) regular use of illicit substances; (iv) history of psychosis, mania, neurological disorder, or neurodevelopmental disorder; and (vi) using cannabis exclusively for medical reasons.

Eligible participants were scheduled to attend an in-person assessment and were asked to abstain from alcohol, cannabis, and other drugs for 48 hours prior to their appointment (verified with self-report and breathalyzer). Upon arrival, participants provided informed consent, and then completed a series of cognitive tasks, semi-structured interviews, and self-report questionnaires that comprised the larger study, including the measures that were the focus of the current analysis (see Measures). Completion of the full assessment took approximately six hours and participants received \$115 CAD for participating.

A total of 167 participants were enrolled, but $n=7$ were found ineligible during the assessment session. Of the $N=160$ eligible participants who completed the study, $n=7$ were missing data on the CUDIT-R, resulting in a final analytic sample of $N=153$. Sample characteristics are in Table 1.

2.2. Measures

2.2.1. Cannabis Use Disorder Identification Test – Revised (CUDIT-R)

The CUDIT-R (Adamson et al., 2010) consists of eight self-report items assessing past-six-month cannabis consumption and cannabis-related behaviours and consequences. Item responses range from 0 to 4 (see Table 2 for item content and response options) and are summed to obtain a total score. The CUDIT-R exhibited adequate internal consistency in our sample ($\alpha=0.77$; $\omega=0.78$).

2.2.2. Structured Clinical Interview for DSM-5 Substance Use Module (SCID-5)

The SCID-5 (First et al., 2015) was used to assess past-year CUD symptoms. Master's-level research assistants administered the SCID-5 with training and ongoing supervision from the study PI (JDW), a licensed clinical psychologist. Diagnostic criteria for CUD were met if two or more symptoms were endorsed in the past year and there was clinically significant distress or impairment as per DSM-5 criteria.

2.2.3. Timeline Follow Back (TLFB)

The calendar-based TLFB interview (Sobell & Sobell, 1992) was used to characterize past-90-day cannabis consumption in the sample. Participants reported whether they had used cannabis each day, and if so, whether they used a flower or non-flower form of cannabis and the grams of cannabis flower consumed if applicable. To aid in reporting, participants were shown a to-scale image depicting different sized cannabis buds and piles of ground cannabis flower, drawn from the Daily Sessions, Frequency, Age of Onset, and Quantity of Cannabis Use Inventory (Cutler & Spradlin, 2017). They also viewed a cannabis substitute (oregano) that was rolled into joints of different weights (Norberg et al., 2012).

2.3. Data analysis

Analyses were conducted in R (R Core Team, 2022) using the *mirt* (Chalmers et al., 2023) and *pROC* (Robin et al., 2021) packages. Code for analyses is available at https://osf.io/2ck9j/?view_only=7273b06e320449aaac69f67c2f1aee7b. We conducted nonparametric receiver operating characteristic (ROC) analyses to examine the ability of the CUDIT-R to distinguish between individuals who met criteria for CUD (“positive” cases) and those who did not (“negative” cases), as measured using the SCID-5. Area under the ROC curve (AUC) was estimated using the trapezoidal rule, with AUC values greater than 0.80 indicating good accuracy in distinguishing between positive and negative cases. We also examined the

sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) of the CUDIT-R at cut-points from 0 to 32. We aimed to identify a cut-point that optimized sensitivity and specificity; however, we prioritized sensitivity given the goal of identifying a cut-point that performed well for detecting potential CUD for further assessment within an SBIRT framework, in which false positives (i.e., referring negative cases for further assessment) may be less of a concern than false negatives (i.e., having positive cases go undetected).

To evaluate the psychometric properties of individual CUDIT-R items, we conducted an item response theory (IRT) analysis. IRT assumes that items of a measure are manifest indicators of a continuous, normally distributed, latent construct (Embretson & Reise, 2013; Harvey & Hammer, 1999). Per IRT, observed responses to an item are jointly determined by the latent construct and the characteristics of the item—specifically, difficulty and discrimination (Embretson & Reise, 2013; Harvey & Hammer, 1999). Given that items of the CUDIT-R are ordinal and polytomous, we fit a graded response model to the data (Samejima, 1997). Overall model fit was considered adequate if the root mean square error of approximation (RMSEA) was <0.06 , and overall model fit and individual item fit were considered adequate if the standardized root mean squared residuals (SRMR) were <0.08 (Hu & Bentler, 1999). In the graded response model, items are described by a slope or discrimination parameter, α , and a category threshold or location parameter, β . There is one discrimination parameter for each item, with higher discrimination parameters indicating that an item is better able to discriminate between individuals with different levels of the latent trait (Baker, 2001; Harvey & Hammer, 1999). There are $k - 1$ location parameters for each item with k response options; each location parameter indicates the level of the standardized latent construct at which the probability of the response lying at or above a given threshold between response options is 0.50 (Baker, 2001;

Harvey & Hammer, 1999). Wider spacing of location parameters suggests an item that is descriptive across a larger range of the latent construct. Relatively lower location parameters suggest lower-difficulty items, whereas relatively higher location parameters suggest higher-difficulty items (Baker, 2001).

3. Results

All participants reported using cannabis in the past six months on the CUDIT-R. The average score on the CUDIT-R was 10.88 ($SD=5.92$), and scores ranged from 2 to 26. Distributions of responses for each CUDIT-R item are displayed in Table 2. Participants endorsed an average of 2.33 ($SD=2.34$) CUD symptoms on the SCID-5, and 53.59% ($n=82$) of participants endorsed two or more CUD symptoms, meeting CUD criteria. Based on symptom count, 30.72% ($n=47$), 11.76% ($n=18$), and 11.11% ($n=17$) of participants were classified as having mild, moderate, and severe CUD, respectively.

On the TLFB, 96.08% ($n=147$) of participants reported using cannabis in the past 90 days, and 22.22% ($n=34$) of participants reported daily or near-daily use. Participants reported using any cannabis on an average of 39.76 ($SD=50.59$) of the past 90 days and using non-flower cannabis products on an average of 5.97 ($SD=15.08$) of the past 90 days. On average, participants used 32.25 ($SD=50.48$) grams of cannabis flower in the past 90 days.

3.1. ROC analyses

The AUC for the CUDIT-R was 0.84 (95% CI=[0.78, 0.91]), suggesting good accuracy in distinguishing between those with and without CUD. Sensitivities, specificities, NPVs, and PPVs from the ROC curve at each cut-point are presented in Table 3. A cut-point of 9 offered 87.80% sensitivity and 70.42% specificity, whereas a cut-point of 10 offered 73.17% sensitivity and 77.46% specificity. Given our prioritization of sensitivity, a cut-point of 9 was favoured.

This cut-point was also associated with an NPV of 83.33% (i.e., 83.33% of those who scored below 9 did not meet for CUD) and a PPV of 77.42% (i.e., 77.42% of those who scored 9 and above did meet for CUD) in our sample. Results of ROC analyses stratified by sex are provided in the supplemental materials for interested readers. However, given that the sample size was modest and stratification by sex leads to fairly small subsamples, the results of the sex-based analyses are considered preliminary and tentative.

3.2. IRT analyses

A graded response model provided adequate fit to the overall CUDIT-R (RMSEA=0.05, SRMR=0.06) and to each CUDIT-R item (all RMSEA<0.06). Item parameters α and β estimated from the graded response model are shown in Table 4. Items 3–6 displayed the highest levels of discrimination ($\alpha=1.85$ – 2.55) and items 2 and 7 displayed the lowest levels of discrimination ($\alpha=0.64$ – 0.67). Location parameters spanned the widest range for item 2 ($\beta=-5.34$ – 4.41), providing information at both very low and very high levels of the latent hazardous use construct. Similarly, items 1 ($\beta=-2.31$ – 0.45) and 6 ($\beta=-0.34$ – 2.33) were descriptive at both below-average and above-average levels of the latent hazardous use construct, although provided narrower coverage relative to item 2. Item 1 exhibited relatively low item difficulty in the sample, whereas item 6 exhibited relatively high item difficulty in the sample. Location parameters for items 3–5 and 7 spanned moderate ranges and were primarily descriptive at above-average levels of the latent hazardous use construct; this was especially true of item 7, for which category thresholds were all located at very high levels of the latent hazardous use construct ($\beta=3.03$ – 6.45), suggesting very high item difficulty in the sample. Location parameters for item 8 suggested relatively low coverage of the latent hazardous use construct, providing

information only at below-average levels ($\beta=-0.89--0.18$) and suggesting relatively low item difficulty in the sample.

4. Discussion

Despite reporting the highest rates of cannabis use and CUD (Government of Canada, 2021; Substance Abuse and Mental Health Services, 2021), young adults are less likely than other age groups to seek treatment for CUD (Wu et al., 2017). Thus, SBIRT approaches may be important for young adults in community settings who engage in regular cannabis use, requiring high-quality CUD screening tools that can identify at-risk young adults to refer for further assessment and treatment. We examined the CUDIT-R as a potential screening tool for young adults who use cannabis, using a relatively diverse sample in terms of sex, racial/ethnic identity, income level, and education level. We extend prior evaluations of the CUDIT-R through our focus on community-based young adults in a jurisdiction in which cannabis is fully legal for non-medical use, and through assessing CUD using a gold standard diagnostic interview.

Results suggest that the CUDIT-R may be effective in identifying young adults with CUD, and that accuracy of the CUDIT-R was similar across male and female sexes. A cut-point of nine exhibited excellent sensitivity (88%) in identifying young adults who meet criteria for CUD. Specificity at this cut-point (70%) was lower, although an over-inclusive cut-point is appropriate to minimize non-detection of potential CUD cases among young adults at risk for CUD. NPV and PPV also supported a cut-point of nine as optimal in minimizing non-detection of potential CUD cases while still avoiding excessive false positives, indicating that 83% of participants who were below this cut-point did not meet SCID-5 criteria for CUD, and 77% of participants who were above this cut-point did meet SCID-5 criteria for CUD.

Consistent with Schultz et al.'s (2019) examination of the CUDIT-R among young adult college students, our recommended cut-point is lower than the cut-point of 13 identified in the original CUDIT-R validation study (Adamson et al., 2010). This difference may be in part attributable to changes in CUD diagnostic criteria from the DSM-IV to DSM-5 or to our use of a community sample of young adults in contrast to Adamson et al.'s (2010) clinical sample of treatment-seeking adults. As young adults perceive cannabis use as less risky relative to other age groups (Waddell, 2022), they may be less likely to recognize and report their cannabis use as hazardous on self-report measures. This may be especially true in jurisdictions in which non-medical cannabis use is fully legal and perceptions of cannabis-related risk are attenuated (Goodman & Hammond, 2022). Thus, a lower cut-point on the self-report CUDIT-R in our young adult sample is unsurprising, and is consistent with lower cut-points for alcohol use disorder screening tools among adolescents and young adults relative to older populations (DeMartini & Carey, 2012; Reinert & Allen, 2007).

In contrast, our recommended cut-point of nine is higher than the cut-point of six recommended by Schultz et al.'s (2019) analysis of young adults, despite our sample's comparable mean CUDIT-R score and similar proportion of participants meeting CUD criteria. This difference likely reflects our use of a more rigorous diagnostic interview, rather than self-report checklist, to determine whether participants in our sample met criteria for CUD. Hence, our higher CUDIT-R cut-point likely reflects a higher threshold for CUD diagnosis when a clinical interview is used to query the clinical significance of symptoms. The higher cut-point we observed relative to Schultz et al. (2019) may also owe to the full legality of non-medical cannabis use in the jurisdiction in which our study took place. In this context, frequent cannabis use may be more normative among non-problem users, and thus the threshold for scores on the

CUDIT-R (which includes items assessing frequency of use and time spent using) that are indicative of CUD may be higher in legal settings.

Individual item properties of the CUDIT-R from IRT analyses also provide insight about the performance of the CUDIT-R in our community-based young adult sample. Consistent with previous findings (Schultz et al., 2019), items 3–6, assessing the inability to stop using cannabis, failure to meet expectations due to cannabis use, time devoted to cannabis use, and memory or concentration problems resulting from cannabis use, were excellent at discriminating between different levels of the latent hazardous use construct. These items were also moderately to highly descriptive, though primarily at above-average levels of the latent hazardous use construct, suggesting relatively high difficulty in the sample.

Items 1 and 8, assessing frequency of use and attempts to reduce use, respectively, were moderately discriminatory and the least difficult. The low difficulty of these items is consistent with previous IRT analyses (Annaheim et al., 2010; Schultz et al., 2019) and indicates that these items do not discriminate among individuals with higher levels of the latent hazardous use construct. Whereas item 8 was similarly discriminatory in our sample and Schultz et al.'s (2019) sample, item 1 was not. Specifically, item 1 showed considerably poorer ability to discriminate between different levels of the latent hazardous use construct in our sample. This may be due to the restriction of our sample to young adults with a history of regular cannabis use, which was not a criterion for inclusion in the Schultz et al. study. In addition, poorer discrimination may reflect elevated rates of high-frequency cannabis use among both problem and non-problem users in our study's legal context.

Items 7 and 2, assessing use with potential for physical hazard and hours per typical cannabis use day spent high, respectively, were the poorest performing items in our sample. Item

7 exhibited poor discrimination and very high difficulty, likely attributable to low-frequency endorsement of the upper response options. Item 2 was similarly poor at discriminating between different levels of the latent hazardous use construct, consistent with a previous IRT analysis that examined this item as part of the longer CUDIT questionnaire (Annaheim et al., 2010). The ability of hours high to discriminate between different levels of hazardous use may be hindered by tolerance. That is, young adults with higher levels of hazardous cannabis use may use more cannabis, but due to higher tolerance, may spend fewer hours high. Consequently, hours high may not differentiate between young adults with different levels of hazardous cannabis use, despite their marked differences in cannabis consumption.

Results of this study should be considered in the context of several limitations. Most notably, our sample size of $N=153$ was modest. However, recent guidance suggests that fewer than 100 participants may be sufficient for ROC analyses of screening tools when the disorder of interest is prevalent and sensitivity is prioritized (Bujang & Adnan, 2016), as was the case in our study. Still, as our sample size was relatively small for polytomous IRT models (e.g., Dai et al., 2021), results should be replicated in larger samples, which could strengthen the generalizability of our results and permit an examination of differential item functioning across relevant demographic groups. Relatedly, as stratifying our sample by sex resulted in modest subsamples, results of sex-stratified ROC analyses should be considered preliminary and tentative. Replication of sex-stratified analyses in larger samples is recommended. In addition, as all participants in the current study were non-treatment-seeking, criteria for moderate and severe CUD were relatively infrequently met. Evaluating the performance of the CUDIT-R in distinguishing between mild, moderate, and severe CUD in larger, treatment-seeking samples may be an important direction for future research.

Several sample characteristics may also have limited the generalizability of our findings. First, our sample had relatively low representation of people reporting current heavy or daily cannabis use, who may be most likely to meet CUD criteria. Second and relatedly, as we required participants to abstain from using cannabis during the 48 hours prior to their assessment, those most at risk of experiencing cannabis withdrawal symptoms may have elected not to participate in the study. Third, all participants reported at least one lifetime period of regular cannabis use. Still, young adults reporting regular cannabis use would be of highest priority for CUD screening. Fourth, we excluded young adults reporting exclusively medical reasons for cannabis use. Fifth, we excluded young adults reporting regular use of illegal substances or severe mental health problems. Sixth, although relatively diverse in terms of sex, racial/ethnic identity, income level, and education level, the current study used a convenience sample. Replicating analyses in future studies with fewer exclusion criteria is needed to bolster the generalizability of our findings.

4.1. Conclusions

In summary, findings support the utility of the CUDIT-R as a sensitive screening tool for detecting CUD among young adults, who are a high priority population for screening efforts. Our findings suggest that a lower cut-point of nine may be optimal when using the CUDIT-R for CUD screening in this age group. Future evaluations of the CUDIT-R in young adult samples may consider removing the item assessing hours high given its poor ability to discriminate between different levels of hazardous use. Alternative indices of consumption, such as the amount of cannabis used in a typical cannabis use day, may also be tested in its place. Confirmatory studies testing a revised version of the CUDIT-R for use among young adults are an important direction for future research.

References

- Adamson, S. J., Kay-Lambkin, F. J., Baker, A. L., Lewin, T. J., Thornton, L., Kelly, B. J., & Sellman, J. D. (2010). An improved brief measure of cannabis misuse: The Cannabis Use Disorders Identification Test-Revised (CUDIT-R). *Drug and Alcohol Dependence*, *110*(1–2), 137–143. <https://doi.org/10.1016/j.drugalcdep.2010.02.017>
- American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition* (Arlington, VA; 5th ed.). American Psychiatric Association.
- Annaheim, B., Scotto, T. J., & Gmel, G. (2010). Revising the Cannabis Use Disorders Identification Test (CUDIT) by means of Item Response Theory. *International Journal of Methods in Psychiatric Research*, *19*(3), 142–155. <https://doi.org/10.1002/mpr.308>
- Baker, F. B. (2001). *The Basics of Item Response Theory. Second Edition*. For full text: <http://ericae>. <https://eric.ed.gov/?id=ED458219>
- Bujang, M. A., & Adnan, T. H. (2016). Requirements for Minimum Sample Size for Sensitivity and Specificity Analysis. *Journal of Clinical and Diagnostic Research : JCDR*, *10*(10), YE01–YE06. <https://doi.org/10.7860/JCDR/2016/18129.8744>
- Chalmers, P., Pritikin, J., Robitzsch, A., Zoltak, M., Kim, K., Falk, C. F., Meade, A., Schneider, L., King, D., Liu, C.-W., & Oguzhan, O. (2023). *mirt: Multidimensional Item Response Theory* (1.38.1) [Computer software]. <https://CRAN.R-project.org/package=mirt>
- Cuttler, C., & Spradlin, A. (2017). Measuring cannabis consumption: Psychometric properties of the Daily Sessions, Frequency, Age of Onset, and Quantity of Cannabis Use Inventory (DFAQ-CU). *PloS One*, *12*(5), e0178194. <https://doi.org/10.1371/journal.pone.0178194>
- Dai, S., Vo, T. T., Kehinde, O. J., He, H., Xue, Y., Demir, C., & Wang, X. (2021). Performance of Polytomous IRT Models With Rating Scale Data: An Investigation Over Sample Size,

- Instrument Length, and Missing Data. *Frontiers in Education*, 6.
<https://www.frontiersin.org/articles/10.3389/feduc.2021.721963>
- DeMartini, K. S., & Carey, K. B. (2012). Optimizing the use of the AUDIT for alcohol screening in college students. *Psychological Assessment*, 24, 954–963.
<https://doi.org/10.1037/a0028519>
- Embretson, S. E., & Reise, S. P. (2013). *Item Response Theory*. Psychology Press.
- First, M. B., Williams, J. B. W., Karg, R. S., & Spitzer, R. L. (2015). *Structured clinical interview for DSM-5—Research version (SCID-5 for DSM-5 Research Version; SCID-5-RV)*. American Psychiatric Association.
- Gette, J. A., Regan, T., & Schumacher, J. A. (2023). Screening, brief intervention, and referral to treatment (SBIRT) for cannabis: A scoping review. *Journal of Substance Use and Addiction Treatment*, 146, 208957. <https://doi.org/10.1016/j.josat.2023.208957>
- Goodman, S., & Hammond, D. (2022). Perceptions of the health risks of cannabis: Estimates from national surveys in Canada and the United States, 2018–2019. *Health Education Research*, 37(2), 61–78. <https://doi.org/10.1093/her/cyac006>
- Goodman, S., Wadsworth, E., Leos-Toro, C., & Hammond, D. (2020). Prevalence and forms of cannabis use in legal vs. Illegal recreational cannabis markets. *International Journal of Drug Policy*, 76, 102658. <https://doi.org/10.1016/j.drugpo.2019.102658>
- Government of Canada. (2021, December 20). *Canadian Alcohol and Drugs Survey (CADS): 2019 Detailed Tables* [Surveys]. <https://www.canada.ca/en/health-canada/services/canadian-alcohol-drugs-survey/2019-summary/detailed-tables.html>

- Gutkind, S., Shmulewitz, D., & Hasin, D. (2023). Sex differences in Cannabis use disorder and associated psychosocial problems among US adults, 2012–2013. *Preventive Medicine, 168*, 107422. <https://doi.org/10.1016/j.ypmed.2023.107422>
- Hammond, D., Wadsworth, E., Reid, J. L., & Burkhalter, R. (2021). Prevalence and modes of cannabis use among youth in Canada, England, and the US, 2017 to 2019. *Drug and Alcohol Dependence, 219*, 108505. <https://doi.org/10.1016/j.drugalcdep.2020.108505>
- Han, B., Compton, W. M., Blanco, C., & Jones, C. M. (2019). Time since first cannabis use and 12-month prevalence of cannabis use disorder among youth and emerging adults in the United States. *Addiction (Abingdon, England), 114*(4), 698–707. <https://doi.org/10.1111/add.14511>
- Harvey, R. J., & Hammer, A. L. (1999). Item response theory. *The Counseling Psychologist, 27*, 353–383. <https://doi.org/10.1177/0011000099273004>
- Hasin, D. S., Kerridge, B. T., Saha, T. D., Huang, B., Pickering, R., Smith, S. M., Jung, J., Zhang, H., & Grant, B. F. (2016). Prevalence and Correlates of DSM-5 Cannabis Use Disorder, 2012-2013: Findings from the National Epidemiologic Survey on Alcohol and Related Conditions-III. *The American Journal of Psychiatry, 173*(6), 588–599. <https://doi.org/10.1176/appi.ajp.2015.15070907>
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*(1), 1–55. <https://doi.org/10.1080/10705519909540118>
- Khan, S. S., Secades-Villa, R., Okuda, M., Wang, S., Pérez-Fuentes, G., Kerridge, B. T., & Blanco, C. (2013). Gender differences in cannabis use disorders: Results from the

- National Epidemiologic Survey of Alcohol and Related Conditions. *Drug and Alcohol Dependence*, 130(1), 101–108. <https://doi.org/10.1016/j.drugalcdep.2012.10.015>
- Kroon, E., Mansueto, A., Kuhns, L., Filbey, F., Wiers, R., & Cousijn, J. (2023). Gender differences in cannabis use disorder symptoms: A network analysis. *Drug and Alcohol Dependence*, 243, 109733. <https://doi.org/10.1016/j.drugalcdep.2022.109733>
- Leung, J., Chan, G. C. K., Hides, L., & Hall, W. D. (2020). What is the prevalence and risk of cannabis use disorders among people who use cannabis? A systematic review and meta-analysis. *Addictive Behaviors*, 109, 106479. <https://doi.org/10.1016/j.addbeh.2020.106479>
- Norberg, M. M., Mackenzie, J., & Copeland, J. (2012). Quantifying cannabis use with the Timeline Followback approach: A psychometric evaluation. *Drug and Alcohol Dependence*, 121(3), 247–252. <https://doi.org/10.1016/j.drugalcdep.2011.09.007>
- R Core Team. (2022). *R: A language and environment for statistical computing* [Computer software]. R Foundation for Statistical Computing. <https://www.R-project.org/>
- Reinert, D. F., & Allen, J. P. (2007). The alcohol use disorders identification test: An update of research findings. *Alcoholism, Clinical and Experimental Research*, 31(2), 185–199. <https://doi.org/10.1111/j.1530-0277.2006.00295.x>
- Risi, M. M., Sokolovsky, A. W., White, H. R., & Jackson, K. M. (2020). Factor Structure of the Cannabis Use Disorders Identification Test Revised (CUDIT-R) for Men and Women. *Cannabis (Albuquerque, N.M.)*, 3(2), 148–156. <https://doi.org/10.26828/cannabis.2020.02.002>
- Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J.-C., Müller, M., code), S. S. (Fast D., Multiclass), M. D. (Hand & T., & CI), Z. B. (DeLong paired test. (2021).

pROC: Display and Analyze ROC Curves (1.18.0) [Computer software].

<https://CRAN.R-project.org/package=pROC>

Samejima, F. (1997). Graded Response Model. In W. J. van der Linden & R. K. Hambleton (Eds.), *Handbook of Modern Item Response Theory* (pp. 85–100). Springer New York.

https://doi.org/10.1007/978-1-4757-2691-6_5

Schultz, N. R., Bassett, D. T., Messina, B. G., & Correia, C. J. (2019). Evaluation of the psychometric properties of the cannabis use disorders identification test—Revised among college students. *Addictive Behaviors, 95*, 11–15.

<https://doi.org/10.1016/j.addbeh.2019.02.016>

Sobell, L. C., & Sobell, M. B. (1992). Timeline follow-back. In *Measuring alcohol consumption* (pp. 41–72). Humana Press.

Substance Abuse and Mental Health Services. (2021). *2020 NSDUH Detailed Tables*. National Survey on Drug Use and Health. <https://www.samhsa.gov/data/report/2020-nsduh-detailed-tables>

Waddell, J. T. (2022). Age-varying time trends in cannabis- and alcohol-related risk perceptions 2002–2019. *Addictive Behaviors, 124*, 107091.

<https://doi.org/10.1016/j.addbeh.2021.107091>

Wu, L.-T., Zhu, H., Mannelli, P., & Swartz, M. S. (2017). Prevalence and correlates of treatment utilization among adults with cannabis use disorder in the United States. *Drug and Alcohol Dependence, 177*, 153–162. <https://doi.org/10.1016/j.drugalcdep.2017.03.037>

Table 1. Sample characteristics

Characteristic	M (<i>SD</i>), range or <i>n</i> (%)
Age	22.21 (2.06), 19–26
Sex	
Male	63 (41.18)
Female	90 (58.82)
Gender ^a	
Man	62 (40.52)
Woman	88 (57.52)
Non-binary	2 (1.31)
Transgender	2 (1.31)
Two-spirit	0 (0.00)
Prefer to specify	1 (0.65)
Sexual Orientation	
Heterosexual/Straight	102 (66.67)
Gay	6 (3.92)
Lesbian	6 (3.92)
Bisexual	32 (20.92)
Asexual	2 (1.31)
Prefer to specify	5 (3.27)
Race/Ethnicity ^a	
White	78 (50.98)
African Descent/Black	27 (17.65)
Asian	26 (16.99)
Pacific Islander	4 (2.61)
Indigenous/Native North American	1 (0.65)
East Indian	16 (10.46)
Hispanic/Latinx	8 (5.23)
Not listed	13 (8.50)
Annual Household Income	
\$19,999 or less	45 (30.61)
\$20,000–\$49,999	36 (24.49)
\$50,000–\$99,999	33 (22.45)
\$100,000–\$199,999	26 (17.69)
\$200,000 or more	7 (4.76)
Prefer not to answer	6 (3.92)
Highest Level of Education	
Less than high school	1 (0.65)
High school diploma or GED	61 (39.87)
Some college	31 (20.26)
Associates degree or technical certificate	7 (4.58)
Bachelor's degree	51 (33.33)
Master's degree or higher	2 (1.31)
SCID-5 number of CUD symptoms	2.33 (2.34), 0–10
CUDIT-R score	10.88 (5.92), 2–26
TLFB used cannabis one or more times in the past 90 days	147 (96.08)
TLFB used cannabis daily/near daily in the past 90 days ^b	34 (22.22)
TLFB number of cannabis use days in the past 90 days	39.76 (30.59), 0–90
TLFB number of non-flower cannabis use days in the past 90 days	5.97 (15.08), 0–90
TLFB grams of cannabis flower used in the past 90 days	32.25 (50.48), 0–226

Note. SCID-5 = Structured Clinical Interview for DSM-5; CUD = cannabis use disorder; CUDIT-R = Cannabis Use Disorder Identification Test – Revised; TLFB = Timeline Follow Back.

^aParticipants could select more than one option and may thus be represented in more than one category.

^bDaily/near daily use was defined as using cannabis on 85% or more of the past 90 days (corresponding to an approximate six or more cannabis use days per week on average).

Table 2. Observed distribution of scores for each item of the CUDIT-R.

Summary of item content	Response				
	0	1	2	3	4
1. Frequency of use ^a	0 (0.00)	19 (12.42)	35 (22.88)	37 (24.18)	62 (40.52)
2. Hours stoned ^b	5 (3.27)	74 (48.37)	52 (33.99)	13 (8.50)	9 (5.88)
3. Unable to stop ^c	92 (60.13)	18 (11.76)	10 (6.54)	18 (11.76)	15 (9.80)
4. Fail to do what is expected ^c	96 (62.75)	34 (22.22)	11 (7.19)	10 (6.54)	2 (1.31)
5. Time devoted ^c	84 (54.90)	33 (21.57)	15 (9.80)	10 (6.54)	11 (7.19)
6. Memory or concentration problems ^c	62 (40.52)	44 (28.76)	23 (15.03)	16 (10.46)	8 (5.23)
7. Physically hazardous ^c	131 (85.62)	14 (9.15)	5 (3.27)	3 (1.96)	0 (0.00)
8. Cutting down ^d	46 (30.07)	–	25 (16.34)	–	82 (53.59)

Note. Values represent the number of participants with percent out of $N = 153$ in parentheses.

^aResponse options: 0 = *never*, 1 = *monthly or less*, 2 = *2-4 times a month*, 3 = *2-3 times a week*, 4 = *4 or more times a week*

^bResponse options: 0 = *less than 1*, 1 = *1 or 2*, 2 = *3 or 4*, 3 = *5 or 6*, 4 = *7 or more*

^cResponse options: 0 = *never*, 1 = *less than monthly*, 2 = *monthly*, 3 = *weekly*, 4 = *daily or almost daily*

^dResponse options: 0 = *never*, 2 = *yes, but not in the past year*, 4 = *yes, during the past year*

Table 3. Specificity and sensitivity for each cut-point of the CUDIT-R for predicting cannabis use disorder

Cut-point	Sensitivity	Specificity	PPV	NPV
0	100.00	0.00	53.59	–
1	100.00	0.00	53.59	–
2	100.00	0.00	53.59	–
3	98.78	2.82	54.00	66.67
4	97.56	9.86	55.56	77.78
5	96.34	23.94	59.40	85.00
6	92.68	36.62	62.81	81.25
7	90.24	46.48	66.07	80.49
8	89.02	59.15	71.57	82.35
9	87.80	70.42	77.42	83.33
10	73.17	77.46	78.95	71.43
11	69.51	83.10	82.61	70.24
12	63.41	88.73	86.67	67.74
13	51.22	90.14	85.71	61.54
14	48.78	94.37	90.91	61.47
15	43.90	97.18	94.74	60.00
16	39.02	98.59	96.97	58.33
17	32.93	100.00	100.00	56.35
18	29.27	100.00	100.00	55.04
19	25.61	100.00	100.00	53.79
20	20.73	100.00	100.00	52.21
21	19.51	100.00	100.00	51.82
22	17.07	100.00	100.00	51.08
23	10.98	100.00	100.00	49.31
24	4.88	100.00	100.00	47.65
25	2.44	100.00	100.00	47.02
26	1.22	100.00	100.00	46.71
27	0.00	100.00	–	46.41
28	0.00	100.00	–	46.41
29	0.00	100.00	–	46.41
30	0.00	100.00	–	46.41
31	0.00	100.00	–	46.41
32	0.00	100.00	–	46.41

Note. CUDIT-R = Cannabis Use Disorder Identification Test Revised; PPV = positive predictive value; NPV = negative predictive value. Bolding indicates the preferred cut-point.

Table 4. CUDIT-R item slope (α) and category threshold (β) parameter estimates from the graded response model

	Slope α	Threshold β_1	Threshold β_2	Threshold β_3	Threshold β_4
1. Frequency of use ^a	0.99	-2.31	-0.76	0.45	–
2. Hours stoned	0.67	-5.34	0.06	2.84	4.41
3. Unable to stop	2.29	0.33	0.77	1.05	1.70
4. Fail to do what is expected	2.20	0.38	1.30	1.86	2.94
5. Time devoted	2.55	0.12	0.83	1.35	1.86
6. Memory or concentration problems	1.85	-0.34	0.66	1.41	2.33
7. Physically hazardous ^a	0.64	3.03	4.84	6.45	–
8. Cutting down ^c	1.29	-0.89	-0.18	–	–

Note. α represents the discrimination parameter for each item, or a slope indicating an item's ability to discriminate between individuals with different levels of the latent hazardous use construct. Each β represents a location parameter, or the point along the latent hazardous use construct at which the probability of a response lying at or above a given category threshold is 0.50.

^aOnly three category thresholds are provided as no participants in the sample provided a response of 0 (*never*).

^bOnly three category thresholds are provided as no participants in the sample provided a response of 4 (*daily or almost daily*).

^cOnly two category thresholds are provided as this item has only three response options.

Supplemental Receiver Operating Characteristic Analyses Stratified by Sex

Given sex differences in the rates and presentation of cannabis use disorder (CUD; Gutkind et al., 2023; Hasin et al., 2016; Khan et al., 2013; Kroon et al., 2023), we repeated all receiver operating characteristic (ROC) analyses in male and female participants separately. However, given the modest sample size, stratification by sex leads to relatively small subsamples of males and female participants; thus, these analyses should be considered preliminary and tentative, and warrant replication in larger samples.

Area under the ROC curve (AUC) suggested good accuracy of the CUDIT-R in distinguishing between those with and without CUD among both male versus female participants (male: AUC = 0.86, 95% CI = [0.77, 0.95]; female: AUC = 0.83, 95% CI = [0.74, 0.92]). The bootstrap method with 2000 samples was used to compare AUCs by sex; AUCs did not differ significantly between male and female participants ($D = -0.47, p = 0.639$). Consistent with ROC analyses in the full sample, sex-stratified sensitivities and specificities favoured a cut-point of 9 in both sexes. This cut-point yielded 85.71% sensitivity and 71.43% specificity in male participants, and 90.00% sensitivity and 70.00% specificity in female participants. This cut-point was also associated with a negative predictive value (NPV) of 71.43% and positive predictive value (PPV) of 85.71% in male participants, and an NPV of 89.74% and PPV of 70.59% in female participants. Sensitivities, specificities, NPVs, and PPVs from the ROC curve at each cut-point stratified by sex are presented in Supplementary Table S1.

References

- Gutkind, S., Shmulewitz, D., & Hasin, D. (2023). Sex differences in Cannabis use disorder and associated psychosocial problems among US adults, 2012–2013. *Preventive Medicine, 168*, 107422. <https://doi.org/10.1016/j.ypmed.2023.107422>
- Hasin, D. S., Kerridge, B. T., Saha, T. D., Huang, B., Pickering, R., Smith, S. M., Jung, J., Zhang, H., & Grant, B. F. (2016). Prevalence and Correlates of DSM-5 Cannabis Use Disorder, 2012-2013: Findings from the National Epidemiologic Survey on Alcohol and Related Conditions-III. *The American Journal of Psychiatry, 173*(6), 588–599. <https://doi.org/10.1176/appi.ajp.2015.15070907>
- Khan, S. S., Secades-Villa, R., Okuda, M., Wang, S., Pérez-Fuentes, G., Kerridge, B. T., & Blanco, C. (2013). Gender differences in cannabis use disorders: Results from the National Epidemiologic Survey of Alcohol and Related Conditions. *Drug and Alcohol Dependence, 130*(1), 101–108. <https://doi.org/10.1016/j.drugalcdep.2012.10.015>
- Kroon, E., Mansueto, A., Kuhns, L., Filbey, F., Wiers, R., & Cousijn, J. (2023). Gender differences in cannabis use disorder symptoms: A network analysis. *Drug and Alcohol Dependence, 243*, 109733. <https://doi.org/10.1016/j.drugalcdep.2022.109733>

Supplementary Table S1. Specificity and sensitivity for each cut-point of the CUDIT-R for predicting cannabis use disorder, stratified by sex

Cut-point	Female participants				Male participants			
	Sensitivity	Specificity	PPV	NPV	Sensitivity	Specificity	PPV	NPV
0	100.00	0.00	44.44	–	100.00	0.00	66.67	–
1	100.00	0.00	44.44	–	100.00	0.00	66.67	–
2	100.00	0.00	44.44	–	100.00	0.00	66.67	–
3	97.50	4.00	44.83	66.67	100.00	0.00	66.67	–
4	95.00	10.00	45.78	71.43	100.00	9.52	68.85	100.00
5	95.00	24.00	50.00	85.71	97.62	23.81	71.93	83.33
6	92.50	40.00	55.22	86.96	92.86	28.57	72.22	66.67
7	90.00	46.00	57.14	85.19	90.48	47.62	77.55	71.43
8	90.00	58.00	63.16	87.88	88.10	61.90	82.22	72.22
9	90.00	70.00	70.59	89.74	85.71	71.43	85.71	71.43
10	75.00	78.00	73.17	79.59	71.43	76.19	85.71	57.14
11	70.00	82.00	75.68	77.36	69.05	85.71	90.63	58.06
12	57.50	88.00	79.31	72.13	69.05	90.48	93.55	59.38
13	45.00	88.00	75.00	66.67	57.14	95.24	96.00	52.63
14	42.50	92.00	80.95	66.67	54.76	100.00	100.00	52.50
15	37.50	96.00	88.24	65.75	50.00	100.00	100.00	50.00
16	32.50	98.00	92.86	64.47	45.24	100.00	100.00	47.73
17	30.00	100.00	100.00	64.10	35.71	100.00	100.00	43.75
18	25.00	100.00	100.00	62.50	33.33	100.00	100.00	42.86
19	20.00	100.00	100.00	60.98	30.95	100.00	100.00	42.00
20	17.50	100.00	100.00	60.24	23.81	100.00	100.00	39.62
21	15.00	100.00	100.00	59.52	23.81	100.00	100.00	39.62
22	12.50	100.00	100.00	58.82	21.43	100.00	100.00	38.89
23	10.00	100.00	100.00	58.14	11.90	100.00	100.00	36.21
24	2.50	100.00	100.00	56.18	7.14	100.00	100.00	35.00
25	0.00	100.00	–	55.56	4.76	100.00	100.00	34.43
26	0.00	100.00	–	55.56	2.38	100.00	100.00	33.87
27	0.00	100.00	–	55.56	0.00	100.00	–	33.33
28	0.00	100.00	–	55.56	0.00	100.00	–	33.33

29	0.00	100.00	–	55.56	0.00	100.00	–	33.33
30	0.00	100.00	–	55.56	0.00	100.00	–	33.33
31	0.00	100.00	–	55.56	0.00	100.00	–	33.33
32	0.00	100.00	–	55.56	0.00	100.00	–	33.33

Note. CUDIT-R = Cannabis Use Disorder Identification Test Revised; PPV = positive predictive value; NPV = negative predictive value. **Bolding indicates the preferred cut-point.**