Drug and Alcohol Dependence Reports 3 (2022) 100043

Contents lists available at ScienceDirect

Drug and Alcohol Dependence Reports

journal homepage: www.elsevier.com/locate/dadr

An evaluation of eight short versions of the Drug Use Disorder Identification Test (DUDIT). A prison population study

Hilde Pape[a,*], Philipp Lobmaier[b,c], Anne Bukten[a,b,d]

[a] University College of Norwegian Correctional Service, P.O. Box 1, 2001 Lillestrøm, Norway
[b] Norwegian Centre for Addiction Research, University of Oslo, Norway
[c] Division of Mental Health and Substance Abuse, Diakonhjemmet Hospital, Oslo, Norway
[d] Section for Clinical Addiction Research, Oslo University Hospital, Norway

ARTICLE INFO

Keywords:
DUDIT
Short versions
Screening
Validation study
Drug problems
Prison population

ABSTRACT

Background: The eleven-item Drug Use Disorder Identification Test (DUDIT) is a recommended screening tool, but its length may impede its use in prison intake assessments. Hence, we examined the performance of eight brief DUDIT screeners against the full DUDIT, employing a sample of male inmates.

Methods: Our study included male participants in the Norwegian Offender Mental Health and Addiction (NorMA) study who reported pre-prison drug use and who had been incarcerated three months or less (n = 251). We performed receiver operating characteristic curve (ROC) analyses and estimated the area under the curve (AUROC) to assess the performance of DUDIT-C (four drug consumption items) and five-item versions that consisted of DUDIT-C and one additional item.

Results: Almost all (95%) screened positive on the full DUDIT (scores ≥6) and 35% had scores that were indicative of drug dependence (scores ≥25). The DUDIT-C performed very well in detecting likely dependence (AUROC=0.950), but some of the five-item versions performed significantly better. Of these, the DUDIT-C + item 5 (craving) had the highest AUROC (0.097). A cut-point of ≥9 on the DUDIT-C and ≥11 on the DUDIT-C + item 5 identified virtually all (98% and 97%, respectively) cases of likely dependence, with a specificity of 73% and 83%, respectively. At these cut-points, the occurrence of false positives was modest (15% and 10%, respectively) and only 4–5% were false negatives.

Conclusions: The DUDIT-C was highly effective in detecting likely drug dependence (according to the full DUDIT), but some combinations of DUDIT-C and one additional item performed better.

1. Introduction

A large proportion of people in prison have a history of illegal drug use (van de Baan et al., 2021; Kouyoumdjian et al., 2014; Stewart, 2009), and drug use disorders are far more prevalent among prison entrants than in the general population (Fazel et al., 2017). The range and severity of harmful outcomes of extensive drug use underscore the importance of effective treatment, and the prison setting may potentially offer a unique opportunity to detect and treat those in need.

Systematic screening is a critically important first step to identify individuals who should be offered drug treatment during incarceration. To what extent such screening occurs is largely unknown, but many European countries screen individuals for alcohol problems shortly after entry into prison (WHO, 2019a). However, validated tools are rarely used (WHO, 2019b), which is probably also the case regarding screening for drug problems. The likely consequence is that the problems often go unnoticed and that many of those in need of treatment never receive professional help.

For such reasons, the correctional service in Norway has been strongly advised to implement universal screening for substance use problems using validated instruments (Oslo Economics and Tyrilistiftelsen, 2020) such as the Drug Use Disorder Identification Test (DUDIT) (Berman et al., 2005). The eleven-item DUDIT was originally intended for use together with the widely recognized Alcohol Use Disorders Identification Test (AUDIT), and the two instruments are parallel in structure, content, and scoring. Moreover, both enable classification of individuals into non-problem users, harmful users, and likely dependent users.

* Corresponding author.
E-mail address: Hilde.Pape@krus.no (H. Pape).

https://doi.org/10.1016/j.dadr.2022.100043
Received 11 January 2022; Received in revised form 9 March 2022; Accepted 10 March 2022

The DUDIT was developed in the early 2000s and is currently freely available in 30 languages.[1] Individuals in the criminal justice system were included in the first validation study of the instrument, which showed that it performed very well in identifying drug dependence according to DSM-4 and ICD-10 criteria (Berman et al., 2005). Several subsequent studies have echoed these results, and a literature review from 2015 concluded that the DUDIT performs "at a level that is at least comparable to, if not higher, than other drug screening instruments" (Hildebrand, 2015, p. 58). More recent research corroborates this conclusion (e.g., Basedow et al., 2021; Klimkiewicz et al., 2020; Sfendla et al., 2017).

According to Klimkiewicz et al. (2020), the DUDIT is currently one of the most frequently recommended instruments for the screening of drug problems. However, the length of the instrument may impede its use in some contexts. The assessment of people who enter prison typically covers a broad range of issues related to health, living conditions, lifestyle, and psychosocial adjustment. Therefore, full versions of screening instruments may be considered too time-consuming. If short versions are unavailable, standardized tools may be deselected in favor of a few unvalidated questions and discretionary assessment of the answers.

1.1. The DUDIT consumption (DUDIT-C) measure

The DUDIT-C consists of four items. They capture the frequency of drug use, the frequency of using more than one drug at the same time, the frequency of drug-taking on a typical day of drug use, and the frequency of being heavily influenced by drugs. The DUDIT-C is analogue to the well-established AUDIT consumption (AUDIT-C) measure, which has been found to detect severe alcohol problems approximately as well as the full AUDIT in various groups and contexts (Kriston et al., 2008; Toner et al., 2019).

The DUDIT-C has been used as a composite drug consumption measure in several studies (e.g., Berman et al., 2015; Bjornestad et al., 2019; Bright et al., 2018). However, only one study (Basedow et al., 2021) has examined its performance as a brief screening tool thus far. It employed a sample of adolescent psychiatric patients with and without a diagnosis of substance use disorder (SUD), and the presence and severity of a drug-related SUD were the main validation standards. The results showed that the complete DUDIT and the DUDIT-C both had outstanding accuracy for detecting SUDs regardless of severity. The complete version performed somewhat better than the DUDIT-C, but the differences were quite small – notably for less severe SUD. Whether the performance of the full and the four-item DUDIT showed statistically significant variation was not reported.

1.2. The present study

In this study, we analyzed data from a prison population sample to investigate the performance of eight brief versions of the DUDIT. Previous analyses of the dataset showed that adding one more item to AUDIT-C significantly improved its effectiveness in detecting alcohol problems (Pape et al., 2021), and we expected this to be the case for the DUDIT-C as well.

In addition to the four-item DUDIT-C, we thus inspected the performance of all possible versions that consisted of the DUDIT-C and one additional item. Moreover, we examined which of these brief screeners that was superior. Finally, we explored cut-points on the DUDIT-C and on the best-performing five-item version.

As in several validation studies of abbreviated AUDIT screeners (e.g., Morojele et al., 2017; Nehlin et al., 2012; Neumann et al., 2012; Pape et al., 2021), we used the respondents' scores on the complete screening tool to specify reference standards. In other words, we applied internal (the full DUDIT) rather than external (e.g., diagnosed drug dependence) validation criteria.

2. Material and methods

We analysed data from the Norwegian Offender Mental Health and Addiction (NorMA) study, which included inmates in almost all (57 of 63) prison units in Norway in 2013/14 (Bukten et al., 2015). There were no pre-defined exclusion criteria, and about 40 percent of the total prison population in Norway took part on the study (n = 1495). The participants responded to a self-report questionnaire that was accessible in Norwegian, English, German, Russian and French.

The causes for non-participation included temporarily absence from the prison, language barriers, and preclusion of study eligibility by prison authorities for security reasons. The sample was, however, representative of the general prison population with respect to many demographic variables. including the percentage of females, the percentage of individuals with a Norwegian citizenship, and country of birth (Bukten et al., 2015; Toresen Lokdam et al., 2021).

Participation was voluntary and based on written informed consent. Confidentiality was ensured and it was pointed out that refraining from participation was not associated with any kind of sanctions.

The NorMA-study was approved by the Norwegian Committee of Research Ethics. Details about data collection and research ethics are reported elsewhere (Bukten et al., 2015).

2.1. Study sample

A majority (56%) of the respondents reported drug use in the year prior to imprisonment (i.e. scores ≥1 on DUDIT item 1). From this group, we selected individuals who had been imprisoned 3 months or less and who had responded to all the 11 DUDIT items – which was the case for 90 percent of those fulfilling the two first inclusion criteria. The number of females in the resulting subsample was low (n = 30), precluding the possibility of performing gender-specific analyses. Therefore, we restricted our analyses to males (n = 251).

2.2. Measures

2.2.1. The Drug Use Disorder Identification Test (DUDIT)

The DUDIT captures use and misuse of illegal substances, excluding alcohol and tobacco use. The original DUDIT consumption items are formulated in present tense without a specified time frame, while the remaining seven items refer to the past year (Berman et al., 2005). In the NorMA-study, all items were reformulated to assess the year before incarceration (Table 1). In the following, we occasionally use the term 'DUDIT-11' when referring to the complete DUDIT.

There are five response options for items 1 to 9 (coded 0, 1, 2, 3, 4) and three for items 10 and 11 (coded 0, 2, 4). The responses are added up and the total score indicates the probability of a drug-related diagnosis. Because we excluded individuals who reported no pre-prison drug use, the total score ranged from 1 (rather than 0) to 44. The DUDIT has generally displayed high internal consistency (Hildebrand, 2015), which was also the case in our sample (Cronbach's α = 0.92).

We followed the DUDIT guidelines (Berman et al., 2005), implying that a positive screen was defined as scores ≥6. Moreover, we made a distinction between harmful drug use (scores 6–24) and likely drug dependence (scores ≥25), which is in accordance with SUD diagnoses in DSM-4. In addition to the DUDIT-C (item 1 + 2 + 3 + 4; scale 1–16, Cronbach's alpha = 0.86), we constructed all possible five-item versions that consisted of the DUDIT-C and one additional item (scale 1–20). Their alpha reliability ranged from 0.81 (DUDIT-C + item 11) to 0.88 (DUDIT-C + item 6).

2.2.2. Pre-prison use of specific drugs

We applied dichotomous measures on frequent use (4+ times a week) of specific drugs (e.g., cannabis, amphetamines, heroin) in the half-

[1] The DUDIT is freely available in 30 languages at https://www.emcdda.europa.eu/drugs-library/drug-use-disorders-identification-test-dudit_en.

Table 1
The eleven DUDIT items as formulated in the present study.

Drug consumption	
1	How often did you use drugs in the last year before incarceration?
2	Did you use more than one type of drug on the same occasion in the last year before incarceration?
3	How many times did you take drugs on a typical day when you used drugs in the last year before incarceration?
4	How many times in the last year before incarceration were you influenced heavily by drugs?
Dependence symptoms and drug-related harm	
5	Have you, in the last year before incarceration, felt that your longing for drugs was so strong that you could not resist it?
6	Has it happened, in the last year before incarceration, that you have not been able to stop taking drugs once you started?
7	How often in the last year before incarceration have you taken drugs and then neglected to do something you should have done?
8	How often in the year before incarceration have you needed to take a drug the morning after heavy drug use the day before?
9	How often in the year before incarceration have you had guilt feelings or a bad conscience because you used drugs?
10[1]	Have you or anyone else been hurt (mentally or physically) because you used drugs?
11[1]	Has a relative or a friend, a doctor or a nurse, or anyone else, been worried about your drug use or said to you that you should stop using drugs?

[1] The time frame for this item was specified in the response options; "No" (scored 0), 'Yes, but not in the year before incarceration' (scored 2), and 'Yes, in the year before incarceration' (scored 4).

year before incarceration. Missing data ranged from 6.8% (cannabis) to 29.5% (heroin).

2.2.3. Other measures

We used data on age (missing data: 6.8%), type of offense (drug-related versus other offences) and current imprisonment length.

2.3. Statistical analyses

We conducted receiver operating characteristic (ROC) analysis (Hanley, 1989), and calculated the area under the curve (AUROC) to examine the performance of the short DUDIT screeners. ROC curves display the true positive rate (sensitivity) against the false positive rate (1-specificity) for all scores, and the AUROC ranges from 0.5 to 1.0. AUROC values between 0.8 and 0.9 are considered "good" and values exceeding 0.90 are "excellent" (Cuparencu et al., 2020). Because we did not use external validation criteria but relied on the full DUDIT to specify reference standards, the likelihood of obtaining high AUROC-values was elevated.

To test whether the different brief versions of the DUDIT performed differently, we compared their AUROCs using z-statistics for paired design (DeLong et al., 1988). We also examined whether statistically significant differences remained significant when adjusting for multiple testing using the Bonferroni method (Bland and Altman, 1995). Furthermore, the combined level of sensitivity and specificity for different cut-points was calculated using Youden's (1950) ($J = \frac{\%\ \text{sensitivity} + \%\ \text{specificity}}{100} - 1$). In addition, we estimated the positive and the negative predictive value of various cut-points.

2.3.1. Sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV)

The higher the sensitivity of a cut-point the lower is its specificity, and vice versa. In our study, sensitivity refers to the percentage of positive cases on the DUDIT-11 that were correctly identified as such by the brief DUDIT screener. Specificity is the percentage of negative cases on the DUDIT-11 that were also negative cases according to the brief screener.

The PPV is the percentage of the positive cases on the brief screener that were also positive cases on the DUDIT-11, while the NPV is the percentage of the negative cases on the brief screener that was also identified as such on the DUDIT-11. Thus, the higher the PPV the lower is the occurrence of false positives, and the higher the NPV the lower is the occurrence of false negatives. In contrast to sensitivity and specificity, the PPV and the NPV both depend on the prevalence of "true" positive cases (as defined by the DUDIT-11) in the sample.

Our view is that sensitivity should be prioritized over specificity when considering cut-points for severe drug problems among people in prison. To identify all or virtually all individuals who probably suffer from drug dependence is imperative as their treatment need is likely substantial. Therefore, we focused solely on cut-points with a sensitivity above 90 percent for identifying likely drug dependence.

Table 2
Descriptive statistics of the study sample.

		%	(n)
Age	18–25 years	29.5	(74)
	26–35 years	36.3	(91)
	≥ 36 years	34.3	(86)
Imprisonment length[1]	< 1 month	41.8	(105)
	1–2 months	33.1	(83)
	> 2 months	25.1	(63)
Type of offense	Drug-related	57.4	(144)
	Other	43.4	(107)
Frequent pre-prison use of specific drugs[2]	Cannabis	52.6	(123)
	Amphetamines	29.3	(61)
	Heroin	23.8	(41)
	Other opioids[3]	18.8	(34)
	GHB	5.9	(11)
DUDIT categories (score range)	Negative screen (1–5)	5.2	(13)
	Harmful drug use (6–24)	35.1	(88)
	Likely dependence (25–44)	59.8	(150)

[1] Those incarcerated > 3 months were excluded.
[2] I.e., 4+ times a week in the half-year before imprisonment.
[3] I.e., opioid-containing medication without perception.

3. Results

3.1. Sample description

Table 2 shows that 70% of the respondents were older than 25 years, 42% had been incarcerated less than a month, and 57% were imprisoned due to drug-related offences. Cannabis was by far the most commonly reported drug that had been used frequently (≥4 times a week) during the half-year before imprisonment, followed by amphetamine, heroin, other opioids, and GHB. Finally, Table 2 shows that 35% had DUDIT scores that were indicative of harmful drug use (scores of 6–24) while 60% had scores in the dependent range (scores ≥25).

In our sample of males who reported drug use in the year before imprisonment, 95% thus screened positive on the complete DUDIT (scores ≥6) and were at least harmful drug users. Hence, a single yes/no-question about pre-prison drug use to *all* male inmates would identify individuals with at least harmful drug use (according to the DUDIT-11) fairly accurately; only 5% of those responding "yes" would be false positives (i.e., scores ≤5), and it goes without saying that the sensitivity

Table 3
AUROCs for detecting likely drug dependence (as classified by DUDIT-11) and p-values for differences between DUDIT-C and five-item versions consisting of DUDIT-C and one additional item.

	AUROC (95% CI)	Comparison with DUDIT-C
DUDIT-C (item 1 + 2 + 3 + 4)	0.950 (0.916- 0.974)	_ _ _
DUDIT-C + item 5	0.979 (0.952–0.993)	<0.001[1]
DUDIT-C + item 6	0.977 (0.950–0.992)	0.001[1]
DUDIT-C + item 7	0.966 (0.936–0.985)	0.010[2]
DUDIT-C + item 8	0.973 (0.944–0.989)	0.001[1]
DUDIT-C + item 9	0.978 (0.951–0.992)	0.010[2]
DUDIT-C + item 10	0.966 (0.916–0.974)	0.070
DUDIT-C + item 11	0.955 (0.922–0.977)	0.419

[1] Bonferroni-adjusted p-value < 0.01 [2] Bonferroni-adjusted p-value >0.05.

would be 100%. Therefore, likely drug dependence (DUDIT-11 scores ≥25) was the sole reference standard in the ROC- analyses.

3.2. Performance of the brief DUDIT screeners and accuracy of cut-points

Table 3 shows that all the brief DUDIT screeners' AUROC for detecting likely drug dependence were well above 0.90. Moreover, all but two of the five-item versions had significantly higher AUROCs than the four-item DUDIT-C. However, when adjusting for multiple testing (Bonferroni correction), only three of these screeners (i.e., DUDIT-C + item 5, 6, and 8, respectively) differed significantly from DUDIT-C alone.

The DUDIT-C + item 5 had the highest AUROC (0.979), and we performed pairwise comparisons to assess whether it differed significantly from any of the other five-item DUDIT screeners in this respect. The results showed that DUDIT-C + item 11 was the only one with a significantly lower AUROC-value ($p = 0.042$), but the Bonferroni-corrected p-value was far above the maximum level of statistical significance ($p>0.9$).

Our results thus showed that a single best-performing short DUDIT screener could not be identified. We still selected the DUDIT-C + item 5 for further analyses – aimed at identifying recommendable cut-points for likely drug dependence. The DUDIT-C alone was also included in these analyses (Table 4).

Regarding the DUDIT-C, a cut-point of ≥5 was the highest whereby *all* individuals with likely drug dependence were detected (i.e., 100% sensitivity and 100% NPV). The specificity was only 37.6%, however, implying that 62.4% of those who did *not* score in the dependent range on DUDIT-11 were misclassified as likely dependent. Moreover, the PPV was 70.4% (i.e., 29.6% false positives). The subsequent cut-points (≥6, ≥7 and ≥8) all had a sensitivity of 99.3%, and we thus focused solely on the cut-point of ≥8 because it (inevitably) had the highest specificity (i.e., 69.3%). Its PPV was 85.0% (15% false positives) and the NPV was 98.6% (i.e., 1.4% false negatives). The cut-point of ≥9 also had a sensitivity that was close to maximum (98.0%), while its specificity was 74.3% (PPV: 86.3%, NPV: 93.5%). The cut-point of ≥10 had the highest J-value, with a sensitivity of 92.4% and a specificity of 82.7% (PPV: 90.3%, NPV: 86.2%).

Moving to the DUDIT-C + item 5, a cut-point of ≥8 was the highest to capture all cases of likely dependence, with a specificity of 62.4% and a PPV of 79.8%. Cut-off scores of ≥9 and ≥10 had identical sensitivity (98.7%). The latter had a specificity of 75.3%, a PPV of 85.5%, and a NPV of 98.4%. The sensitivity was also high for cut-points of ≥11 (97.3%) and ≥12 (96.0%), as was their specificity (83.2% and 88.1%, respectively), PPV (89.6% and 92.3%, respectively) and NPV (95.5% and 93.7%, respectively). A cut-point of ≥13 yielded the highest J-value, with a sensitivity of 94.5% and a specificity of 90.9% (PPV: 90.3%, NPV: 86.2%).

4. Discussion

Non-problematic drug use was very rare in our study of males who reported any use of illegal substances n the year before imprisonment. Specifically, almost all (95%) had a DUDIT positive screen, which is indicative of at least harmful drug use, and six in ten had scores in the dependent range. Precisely because of the extremely high prevalence of the DUDIT positive screens, likely drug dependence (as classified by the full DUDIT) was the one and only reference standard in the ROC-analyses.

We examined the performance of the four-item DUDIT-C as well as all possible five-item versions that consisted of DUDIT-C and one additional item. The results showed that all these brief screeners performed very well, yielding AUROCs well above 0.90. Some of the five-item versions performed significantly better than DUDIT-C, but a single best-performing brief screener could not be identified. However, DUDIT-C + item 5 (craving) had a slightly higher AUROC than the other five-item DUDIT screeners.

Our validation study differed from that of Basedow et al. (2021) in several respects. They evaluated the DUDIT-C and employed a clinical sample of mid-teen boys and girls with mental health and/or drug use problems. In contrast, our prison population sample was restricted to males who reported pre-prison drug use and a solid majority (70%) were older than 25 years of age. Moreover, we applied an internal reference standard (i.e., likely drug dependence according to the full DUDIT) while Basedow et al. (2021) examined the DUDIT-C against DSM-5 criteria of drug-related SUDs – which was an important strength of their study. It is worth noting that both studies found that the DUDIT-C was highly effective in detecting severe drug problems.

4.1. Cut-points: results and considerations

When considering cut-off scores on the brief DUDIT screeners, one should keep in mind that the standard cut-offs on the full DUDIT are based on DSM-4 (and not DSM-5). We inspected cut-points on the DUDIT-C and the DUDIT-C + item 5 and found that the highest threshold level that identified *all* individuals with likely drug dependence (according to the full DUDIT) was ≥5 and ≥8, respectively. The cut-points with the highest Youden's J were ≥10 (DUDIT-C) and ≥13 (DUDIT-C + item 5), and both met our requirement of a sensitivity above 90 percent.

A data-driven approach of detecting cut-points typically relies on Youden's J, without pre-defined priorities regarding sensitivity or specificity. This approach seems reasonable when the purpose is to estimate the prevalence of drug problems in a population and the resources needed to meet them. However, if the aim is to identify individuals in likely need of interventions because their drug use poses grave threats to health and psychosocial functioning, minimizing false negatives is essential.

On the other hand, choosing cut-off scores that minimize or eliminate false negative cases may result in a high percentage of false positives. This, in turn, may bring about increased costs in terms of unnecessary further assessment, diagnostic evaluation, and maybe also delivery of more intensive interventions than required. In many countries, there is a paucity of treatment services for inmates with substance use problems (Belenko et al., 2013; Graham et al., 2012; Pape et al., 2020), which is another issue of potential relevance in this respect.

If the aim of using a brief DUDIT screener is to identify *all* individuals who are likely drug dependent, our results indicated that the DUDIT-C was suboptimal. The cut-point in question (scores of ≥5) had very low specificity (38%); hence, the percentage of false positives was high (31%). DUDIT-C + item 5 performed better; the highest cut-point that detected all cases of likely dependence (scores of ≥8) had much higher specificity (62%) and a lower occurrence of false positive cases (20%).

If identification of *almost* all (rather than *absolutely* all) individuals with likely drug dependence is considered acceptable, the overall accuracy of the selected cut-points will inevitably be higher. Our study

Table 4
Sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) of selected cut-points on DUDIT-C and DUDIT-C + item 5 in identifying likely drug dependence (as classified by DUDIT-11). Optimal cut-points according to Youden's J in bold.

DUDIT-C						DUDIT-C + item 5					
Cut- point[1]	Sens.%	Spec.%	*J*	PPV%	NPV%	Cut- point[2]	Sens.%	Spec.%	*J*	PPV%	NPV%
≥5	100	37.6	0.38	70.4	100	≥8	100	62.4	0.62	79.8	100
≥8	99.3	69.3	0.69	82.8	98.6	≥10	98.7	75.3	0.74	85.5	98.4
≥9	98.0	74.3	0.72	85.0	96.2	≥11	97.3	83.2	0.80	89.6	95.5
≥10	**94.0**	**82.2**	**0.76**	88.7	90.2	≥12	96.0	88.1	0.83	92.3	93.7
—	—	—	—	—	—	**≥13**	**93.3**	**94.1**	**0.87**	95.9	90.5

[1] Cut-points of ≥ 6 and ≥7 had identical sensitivity as the cut-point of ≥8 and are therefore not displayed.
[2] The cut-point of ≥ 9 had identical sensitivity as the cut-point of ≥10 and is therefore not displayed.

showed that a cut-point of ≥9 on the DUDIT-C and ≥11 on the DUDIT-C + item 5 detected 98% and 97% of the "true" positive cases, respectively. Choosing these cut-points rather than those with 100% sensitivity would substantially increase the specificity (to 73% and 83%, respectively) and reduce the problem of false positives (to 15% and 10%, respectively). The percentage of false negative cases at these cut-points was low for both the DUDIT-C (4%) and the DUDIT-C + item 5 (5%).

The sensitivity of the cut-point with the highest Youden's J on each of the two brief screeners (scores ≥10 on DUDIT-C and scores ≥13 on DUDIT-C + item 5) barely varied (94% and 93%, respectively). However, the specificity was markedly lower for the DUDIT-C (82%) than for the DUDIT-C + item 5 (94%), and the occurrence of false positives varied accordingly (11% and 4%, respectively). Thus, our study indicated that choosing these cut-points is not an option unless a rather high percentage of false negatives is considered acceptable. Specifically, 10 percent of those scooring below the cut-points in question on either the DUDIT-C or the DUDIT-C + item 5 had scores in the dependent range on the full DUDIT.

One should keep in mind that the magnitude of misclassifications depends on the prevalence of "true" positive cases among those being screened. The percentage of false positive and false negative cases of likely drug dependence for various cut-points on the brief DUDIT screeners may thus show substantial variation across population subgroups.

4.2. Limitations and methodological considerations

We relied solely on the full DUDIT to specify the reference standard, which increased the likelihood of obtaining high AUROCs. This is an important limitation of our study. The lack of external validation criteria (e.g., drug-related diagnoses) also precluded the possibility to test whether the performance of the brief DUDIT screeners differed significantly from the performance of the complete DUDIT.

Moreover, all the DUDIT-items in our study were reformulated to assess illegal drug use and related harm in the year before imprisonment, and the validity of the responses may have been hampered by inaccurate recall. To reduce recall bias, we excluded individuals who had been imprisoned for more than three months. Of those included in our analyses, a sizable proportion (42%) had spent less than a month in prison. Nevertheless, it would have been advantageous if *all* the respondents had been recruited shortly after entry into prison.

The original DUDIT comes with a list of specified drugs and psychoactive medications (non-prescribed), which was not included in the NorMA-questionnaire. Consequently, the respondents responded to the items without reference to the specific substance(s) that they had used. The drug that was the main source of their problems and dependence symptoms is therefore unknown.

Another limitation is that the brief DUDIT screeners were embedded in the full DUDIT, which might have affected the responses to the items in question. Moreover, our study was restricted to males and the results may well be different for females.

Finally, one should take into consideration that in practice, questions about drug use are typically being asked by prison officers in a face-to-face setting. Disclosing one's pre-prison heavy drug use and dependence symptoms in this context may potentially entail more invasive drug searches during incarceration and more restrictive detention conditions (Malloch, 2001; Kolind and Duke, 2016). Hence, individuals are more likely to underreport to their drug problems when in a "natural" prison setting than when they voluntarily respond to a self-administered questionnaire in a confidential research context.

4.3. Implications and suggestions for future research

Validated tools for screening drug use problems, such as the DUDIT, are crucial for identifying likely treatment needs. However, the length of the DUDIT may limit its use in the ordinary assessment of people entering prison. Shorter alternatives are more likely to be put into practice, which highlights the potential importance of our study.

No previous study has examined the performance of various brief versions of the DUDIT. Our results were promising, but more research is required to draw reasonably firm conclusions. Future validation studies should employ external reference standards and investigate the effectiveness of brief DUDIT screeners in both prison populations and other population groups.

The performance of a brief DUDIT screener as well as its recommendable cut-points may differ for males and females, and such gender differences may vary across population groups. Thus, drug-related SUDs are generally far more prevalent among males than females, yet the opposite has been found in many prison population studies (Binswanger et al., 2010; Fazel et al., 2017). Hence, future validation studies should include both genders, and analyze males and females separately.

5. Conclusions

The four-item DUDIT-C was highly effective in identifying individuals with likely drug dependence (as classified by the full DUDIT), but some combinations of the DUDIT-C and one additional item performed even better.

Declaration of Competing Interest

None of the authors have any conflict of interest to declare.

CRediT authorship contribution statement

Hilde Pape: Formal analysis, Investigation, Writing – original draft, Writing – review & editing. **Philipp Lobmaier:** Visualization, Writing – original draft, Investigation, Writing – review & editing. **Anne Bukten:** Investigation, Writing – original draft, Data curation, Formal analysis, Writing – review & editing.

Acknowledgements

The NorMA-study was funded by the South-Eastern Norway Regional Health Authority (grant number: 2019091) and The Norwegian center for Addiction Research (SERAF) at the University of Oslo. The authors wish to thank the individuals who participated in the study, and the prison management and the staff members at the local prisons for assisting in collecting data.

Author disclosure

All authors have read and approved the manuscript. The article is our original work. It has not received prior publication and is not currently under consideration for publication elsewhere.

References

Basedow, L.A., Kuitunen-Paul, S., Eichler, A., Roessner, V., Golub, Y., 2021. Diagnostic accuracy of the drug use disorder identification Test DUDIT and its short form, the DUDIT-C, in German adolescent psychiatric patients. Front. Psychol. 12, 678819.

Belenko, S., Hiller, M., Hamilton, L., 2013. Treating substance use disorders in the criminal justice system. Curr. Psychiatry Rep. 15 (11), 414.

Berman, A., Bergman, H., Palmstierna, T., Schlyter, F., 2005. DUDIT Manual the Drug Use Disorders Identification Test. Karolinska Institutet, Department of Clinical Neuroscience, London.

Berman, A., Wennberg, P., Sinadinovic, K., 2015. Changes in mental and physical well-being among problematic alcohol and drug users in 12-month Internet-based intervention trials. Psychol. Addict. Behav. 29 (1), 97–105.

Binswanger, I.A., Merrill, J.O., Krueger, P.M., White, M.C., Booth, R.E., et al., 2010. Gender differences in chronic medical, psychiatric, and substance-dependence disorders among jail inmates. Am. J. Public Health 100 (3), 476–482.

Bjornestad, J., Svendsen, T.S., Slyngstad, T.E., Erga, A.H., McKay, J.R., et al., 2019. A life more ordinary" processes of 5-year recovery from substance abuse. Experiences of 30 recovered service users. Front. Psychol. 10, 689.

Bland, J.M., Altman, D.G., 1995. Multiple significance tests: the Bonferroni method. BMJ 310 (6973), 170.

Bright, S., Walsh, K., Williams, C., 2018. Point prevalence and patterns of mental health comorbidity among people accessing Australia's first older adult–specific alcohol and other drug treatment service. J. Dual Diagn. 14 (1), 70–75.

Bukten, A., Lund, I.O., Rognli, E.B., Stavseth, M.R., Lobmaier, P., et al., 2015. The norwegian offender mental health and addiction study. Design and implementation of a national survey and prospective cohort study. Subst. Abuse 9 (Supplement 2), 59–66.

Cuparencu, C., Rinnan, Å., Silvestre, M.P., Poppitt, S.D., Raben, A., et al., 2020. The anserine to carnosine ratio: an excellent discriminator between white and red meats consumed by free-living overweight participants of the PREVIEW study. Eur. J. Nutr. 60, 179–192.

DeLong, E.R., DeLong, D.M., Clarke-Pearson, D.L., 1988. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 837–845.

Fazel, S., Yoon, I.A., Hayes, A.J., 2017. Substance use disorders in prisoners: an updated systematic review and meta-regression analysis in recently incarcerated men and women. Addiction 112 (10), 1725–1739.

Graham, L., Parkes, T., McAuley, A., Doi, L., 2012. Alcohol Problems in the Criminal Justice System: an Opportunity for Intervention. World Health Organisation, Denmark.

Hanley, J.A., 1989. Receiver operating characteristic ROC methodology: the state of the art. Crit. Rev. Diagn. Imaging 29 (3), 307–335.

Hildebrand, M., 2015. The psychometric properties of the drug use disorders identification test DUDIT: a review of recent research. J. Subs. Abuse Treat. 53, 52–59.

Klimkiewicz, A., Jakubczyk, A., Mach, A., Abramowska, M., Serafin, P., et al., 2020. Psychometric properties of the polish version of the drug-use disorders identification test. Eur. Addict. Res. 26 (3), 131–140.

Kolind, T., Duke, K., 2016. Drugs in prisons: exploring use, control, treatment and policy. Drugs Educ. Prev. Policy 23 (2), 89–92.

Kouyoumdjian, F.G., Calzavara, L.M., Kiefer, L., Main, C., Bondy, S.J., 2014. Drug use prior to incarceration and associated socio-behavioural factors among males in a provincial correctional facility in Ontario, Canada. Can. J. Public Health 105 (3), e198–e202.

Kriston, L., Hölzel, L., Weiser, A.K., Berner, M.M., Härter, M., 2008. Meta-analysis: are 3 questions enough to detect unhealthy alcohol use? Ann. Intern. Med. 149 (12), 879–888.

Malloch, M., 2001. Women, Drugs and custody: The experiences of Women Drug Users in Prison (chapter 5). Waterside Press, Winchester, UK.

Morojele, N.K., Nkosi, S., Kekwaletswe, C.T., Shuper, P.A., Manda, S.O., et al., 2017. Utility of brief versions of the alcohol use disorders identification test AUDIT to identify excessive drinking among patients in HIV care in South Africa. J. Stud. Alcohol Drugs 78 (1), 88–96.

Nehlin, C., Fredriksson, A., Jansson, L., 2012. Brief alcohol screening in a clinical psychiatric population: special attention needed. Drug Alcohol Rev. 31 (4), 538–543.

Neumann, T., Linnen, H., Kip, M., Grittner, U., Weiβ-Gerlach, E., et al., 2012. Does the alcohol use disorders identification test–consumption identify the same patient population as the full 10-item alcohol use disorders identification test? J. Subst. Abuse Treat. 43 (1), 80–85.

Oslo Economics and Tyrilistiftelsen, 2020. Vurdering av det samlede tilbudet under straffegjennomføring for personer med rusmiddelproblematikk. (A report to the Norwegian Ministry of Justice and Public Security and the Norwegian Ministry of Health and Care Services). Oslo Economics, Oslo. Available at https://www.regjeringen.no/contentassets/1ac47ba1f08b42caa6612fda2c17ddf0/straffegjennomforing-og-rusmiddelproblematikk.pdf.

Pape, H., Rossow, I., Bukten, A., 2020. Alcohol problems among prisoners: subgroup variations, concurrent drug problems, and treatment needs. Eur. Add. Res. 1–10.

Pape, H., Rossow, I., Bukten, A., 2021. Are short AUDIT screeners effective in identifying unhealthy drinking of varying severity? A prison population study. Drug Alcohol Depend, 109153.

Sfendla, A., Zouini, B., Lemrani, D., Berman, A.H., Senhaji, M., et al., 2017. Psychometric properties of the Arabic version of the drug use disorders identification test DUDIT in clinical, prison inmate, and student samples. Int. J. Behav. Med. 24 (2), 280–287.

Stewart, D., 2009. Drug use and perceived treatment need among newly sentenced prisoners in England and wales. Addiction 104 (2), 243–247.

Toner, P., Böhnke, J.R., Andersen, P., McCambridge, J., 2019. Alcohol screening and assessment measures for young people: a systematic review and meta-analysis of validation studies. Drug Alcohol Depend. (202) 39–49.

Toresen Lokdam, N., Riksheim Stavseth, M., Bukten, A., 2021. Exploring the external validity of survey data with triangulation: a case study from the Norwegian offender mental health and addiction NorMA study. Res. Methods Med. Health Sci. doi:10.1177/26320843211061298.

van de Baan, F.C., Montanari, L., Royuela, L., Lemmens, P.H., 2021. Prevalence of illicit drug use before imprisonment in Europe: results from a comprehensive literature review. Drugs: Educ. Prev. Policy 1–12. https://www.tandfonline.com/doi/pdf/10.1080/09687637.2021.1879022.

WHO, 2019a. Health in Prisons: Fact Sheet For 38 European Countries. WHO Regional office for Europe, Copenhagen 2019. Available at: http://www.euro.who.int/en/health-topics/health-determinants/prisons-and-health/publications/2019/health-in-prisons-fact-sheets-for-38-european-countries-2019 .

WHO, 2019b. Screening for Harmful Use of Alcohol. Health in prisons European Database Available at: http://apps.who.int/gho/data/node.prisons.Screening_Harmful_Use_Alcohol?lang=en2019 .

Youden, W.J., 1950. Index for rating diagnostic tests. Cancer 3 (1), 32–35.