Development of a Routine Outcome Monitoring Instrument for Use with Clients in the New Zealand Alcohol and Other Drug Treatment Sector: the Alcohol and Drug Outcome Measure (ADOM)

Justin Pulford, *CRRC, Waitemata DHB*
Daryle E Deering, *NAC, Christchurch School of Medicine and Health Sciences*
Gail Robinson, *CADS, Waitemata DHB*
Amanda Wheeler, *CRRC, Waitemata DHB*
Simon J Adamson, *NAC, Christchurch School of Medicine and Health Sciences*
Christopher MA Frampton, *Christchurch School of Medicine and Health Sciences*
Lucy Dunbar, *Christchurch School of Medicine and Health Sciences*
Grant Paton-Simpson, *Paton-Simpson Associates Limited, Auckland*

This paper describes the development and evaluation of the Alcohol and Drug Outcome Measure (ADOM), a brief 18-item, two part, outcome monitoring instrument designed for routine use with clients in the New Zealand Alcohol and Other Drug (AOD) treatment sector. The development of the ADOM was informed by an expert panel of AOD clinicians (*n* = 6), key informant interviews (*n* = 14) and a pilot with AOD treatment clients and their respective clinicians (*n* = 25). The psychometric properties of the ADOM were tested with clients (*n* = 63) across a range of AOD treatment services. Testing involved clinicians administering the ADOM at treatment admission, one-to-seven days post-admission, and four-to-six weeks post-admission. Analyses of the test-retest reliability, concurrent validity and sensitivity to change of Part A of the ADOM, covering type and frequency of substance use, consistently produced excellent results. Comparable results for Part B of the instrument, covering associated psychosocial issues, were generally satisfactory. The ADOM has the potential to be used as a core AOD outcomes monitoring instrument.

Internationally, there is an increasing call for the implementation of routine outcome monitoring during alcohol and other drug (AOD) treatment provision (Lawrinson, Roche, & Copeland, 2009; McLellan, Mckay, Forman, Cacciola, & Kemp, 2005). The routine collection of client outcome data provides the means to monitor clients' treatment progress over time, informs individualised clinical decision making, and actively involves the client in treatment review and evaluation processes (Deering, Sellman, Adamson, Horn, & Frampton, 2008); the ultimate aim being to improve the quality of care in everyday clinical settings. Accordingly, a number of outcome monitoring instruments have recently been developed for use in specific jurisdictions or AOD service types (Lawrinson, Copeland, & Indig, 2005; Marsden et al., 2008; Simpson, Lawrinson, Copeland, & Gates, 2009).

In the New Zealand AOD treatment sector there is considerable interest in routine outcome monitoring. Outcome monitoring instruments suited for use in New Zealand local youth AOD treatment services and methadone maintenance services have been developed (Christie et al., 2007; Deering et al., 2008) and the Ministry of Health has invested substantial resource in outcome monitoring through initiatives such as the Mental Health Research and Development Strategy (MHRDS). Through this strategy a team of researchers were commissioned to define a routine outcome monitoring system of potential benefit to AOD clinicians and clients within the context of day-to-day treatment (Deering et al., 2004).

This initial work incorporated a survey of AOD services which identified a strong endorsement for the value of routine outcome measurement utilising brief and multidimensional AOD instruments suited to the New Zealand setting. Further consultation with clinicians and consumers identified that such an instrument would be completed collaboratively between client and clinician, would inform immediate clinical decision-making and would allow client progress to be monitored over time. Based on feedback obtained during the course of this preliminary phase, as well as previous comment in the AOD literature (McLellan et al., 2005; Teesson, Clement, Copeland, Conroy, & Reid, 2000), it was recommended that in order to be feasible and clinically useful in day to day practice the instrument needed to be acceptable, relevant and of value to clinicians and clients, brief (5-10 minutes), multi-dimensional, easy to administer and interpret,

psychometrically sound, and sensitive to change.

In response to this recommendation, the MHRDS (now known as Te Pou) provided additional funding to develop the proposed outcome monitoring instrument. This paper describes the development, test-retest reliability, concurrent validity and sensitivity to change of the resulting instrument: the Alcohol and Drug Outcome Measure (ADOM).

METHOD

Development and design

An ADOM prototype was presented in the originally commissioned report (Deering et al., 2004). The proposed prototype, however, had not been subjected to critical review or presented for stakeholder feedback. Accordingly, the design methodology in the present study was based on refining the original ADOM prototype via a series of consultation and review processes. This included, in sequential order, consultation with: an expert panel of six senior AOD treatment clinicians selected on the basis that they were recognised clinical leaders in the AOD treatment sector and had an understanding of outcome monitoring issues; 14 key informants representing AOD service managers, clinicians and clients, including Māori and Pacific Island representatives, from a range of different services across New Zealand; and 25 AOD treatment clients and their respective clinicians.

Participants at each stage of consultation were requested to critique suggested questions on the basis of their face validity and perceived utility in an AOD clinical context and/or to suggest alternative possible questions. This emphasis was consistent with the aim of developing an outcome monitoring instrument primarily suited to informing clinical decision making at the clinician/client level. Amendments were made to the ADOM prototype following each round of consultation. All amendments were determined in consultation with a project advisory board comprising AOD treatment workforce, Māori, Pacific, consumer, and Ministry of Health representatives as well as an independent contractor experienced in outcome monitoring and health sector information technology systems. This process resulted in an 18-item questionnaire split into two discrete sections: Section A covering type and frequency of substance use (11 items) and Section B covering associated psychosocial issues (7 items). A copy of the ADOM is presented as an Appendix.

Psychometric testing

The aims of the psychometric testing stage were to assess the test-retest reliability, concurrent validity and sensitivity to change of the ADOM within the available funding resource. In order to ensure adequate statistical power for the respective analyses (described below), at least 50 new clients (admissions) to the participating AOD treatment services were required to complete the ADOM at three time points: admission, one-to-seven days post-admission and four-to-six weeks post-admission. In addition, the participating clients were required to complete a suite of comparison measures at the first and third assessment points.

Study setting

The study took place in seven general outpatient AOD treatment services and two opioid substitution treatment services located in the urban centres of Auckland and Christchurch, New Zealand. Collectively, these services are the primary providers of specialist AOD treatment in their respective regions. All participating services employ clinicians from a range of professional backgrounds and operate according to a harm reduction approach.

Ethical approval was obtained for this project from a Ministry of Health, Health and Disability ethics committee.

Procedure and recruitment

Participants were recruited via the referral and allocation procedures of the involved clinical service. Following training in research procedures and protocols and within their service context, AOD clinicians informed their allocated clients at treatment entry about the study and gained written informed consent from those who volunteered to participate.

At the initial interview, participants were asked to complete three different substance use questionnaires and two questionnaires on health and functioning in person, in collaboration with their clinician (described below). Each of the questionnaires was to be completed as a distinct task, with as little cross-questionnaire influence on answers as possible. Within seven days, participating clients completed the ADOM only, either in person with their clinician or over the telephone. Four to six weeks following the initial interview, participants again completed the five questionnaires in person with their clinician. Individual feedback was provided on treatment progress based on questionnaire responses.

All participants were provided with a $20 petrol voucher at the completion of data collection to contribute to travel costs.

Comparative measures

In addition to the ADOM, the measures listed below were administered to participating clients at the first (treatment admission) and third (four to six weeks post admission) assessment points.

- Degree of Drug use Index (DDI): A nine-item questionnaire (with demonstrated validity in the New Zealand clinical setting for Māori and non-Māori clients), designed to assess type, frequency and, in the case of alcohol, quantity of alcohol, and other drug use over the past four weeks, including frequency of any injecting drug use (Deering et al., 2008).
- Timeline Follow Back (TLFB): The TLFB is a calendar-based method for assessing type, frequency and quantity of AOD use over a specified period. In this study, a reference period of 28 days was employed (Sobell & Sobell, 1996).
- Sections two to four of the Treatment Outcome Profile (TOP): Thirteen yes/no, scale (0-20) or numeric (0-28) response questions that collectively examine injecting risk behaviour, criminal activity, and health and social functioning over a four week period (Marsden et al., 2008).
- Questions four and five of the SF-36 Health Survey (SF-36): Two multiple response questions, collectively

comprising seven 'yes/no' items, that examine whether the client has experienced any number of specified problems with their work or other regular daily activities as a result of their physical health or emotional problems (Medical Outcomes Trust, 1994; Ware & Sherbourne, 1992).

The DDI and TLFB were employed as comparative measures for Part A of the ADOM, whilst the TOP and SF-36 questions were employed as comparative measures for Part B.

Analysis

Test-retest reliability was assessed using Cohen's Kappa (k) for categorical and ordinal data and intraclass correlation coefficients (ICC) for continuous data. Concurrent validity was assessed using Cohen's Kappa for categorical data and Spearman's rank correlation coefficient (*r*) for continuous data. Sensitivity to change was assessed using paired sample t-tests and the Reliable Change Index (RCI) for continuous data, the McNemar test for categorical data and the Wilcoxin Signed Ranks test for ordinal data.

RESULTS

Participation

A total of 63 AOD treatment clients successfully completed the baseline interview, 61 of whom completed the test-retest interview (although six of these fell outside of the one to seven day retest period) and 56 of whom completed the sensitivity to change interview. Participant characteristics are presented in Table 1. As can be seen, despite a low level of participant dropout (11%), the characteristics of the sample remain consistent across each interview period.

Table 1: Participant characteristics across the three interview points‡

Characteristic	Sample		
	Baseline	Test-Retest	Change Senstivity
	(*n* = 63)	(*n* = 61)	(*n* = 56)
Gender: *n* (%)			
Male	37 (58.7)	35 (57.4)	30 (53.6)
Female	26 (41.3)	26 (42.6)	26 (46.4)
Ethnicity: *n* (%)			
NZ European	39 (61.9)	37 (60.7)	35 (62.5)
Māori	11 (17.5)	11 (18.0)	10 (17.9)
Pacific Peoples	7 (11.1)	7 (11.4)	7 (12.5)
Other	6 (9.5)	6 (9.8)	4 (7.1)
Substance use at baseline: *n* (%) §			
Alcohol	45 (72.6)	45 (75.0)	41 (74.5)
Cannabis	31 (49.2)	31 (50.8)	30 (53.6)
Amphetamine	13 (21.0)	13 (21.7)	11 (20.0)
Opioids	23 (36.5)	22 (36.1)	21 (37.5)
Sedative/ tranquiliser	9 (14.3)	9 (14.8)	8 (14.3)
Other drug	3 (4.8)	3 (4.9)	1 (1.8)
Cigarettes/ Nicotine	50 (82.0)	48 (81.4)	45 (83.3)
Primary substance use at baseline: *n* (%)¥			
Alcohol	27 (46.6)	26 (46.4)	23 (45.1)
Cannabis	5 (8.6)	5 (8.9)	5 (9.8)
Amphetamine	6 (10.3)	6 (10.7)	5 (9.8)
Opioids	22 (37.9)	21 (37.5)	20 (39.2)
Sedative/ tranquiliser	1 (1.7)	1 (1.8)	1 (2.0)
Other drug	0 (0)	0 (0)	0 (0)
Cigarettes/ Nicotine	6 (10.3)	5 (8.9)	5 (9.8)

‡ Listed frequencies and percentages based on valid data available for each measure, which may not always equal the stated 'n'.
§ Based on past 28 days.
¥ Defined as main substance of concern (participants could identify more than one).

Test-Retest Reliability

Parts A and B of the pilot ADOM were completed by participants after a mean of 3.1 days (*SD* = 2.1)[1] .

ADOM Part A

The results of test-retest analyses for Part A of the ADOM are shown in Table 2. ICCs for the continuous 'mean days used' and 'mean units on typical day' measures were above 0.75 in every case, indicating excellent test-retest reliability (Fleiss, 1991). The ICC statistic was not computed for the 'other drugs - mean days use' measure (question seven) due to low sample size on this variable (only three participants reported 'other drug' use). For the categorical (yes/no) measures of 'primary substance of concern' and 'shared injecting equipment, a *k* of 0.60 or higher indicates good test-retest reliability, and a k of 0.40 may be considered a minimum acceptable value (Landis & Koch, 1977). Thus, test-retest reliability was very good for every categorical measure with the exception of 'primary substance of concern – cigarettes/nicotine' (0.49) and 'shared injecting equipment' (0.49).

ADOM Part B

Kappa scores for each question listed in Part B of the ADOM are presented in Table 2, and are all acceptable, ranging from 0.45 (family/friends conflict) to 0.60 (engaged in work/other activity).

Concurrent Validity

ADOM Part A

Spearman's rank correlation coefficients between comparable parts of ADOM Part A, the DDI and the TLFB were calculated based on data obtained at the baseline interview. These data are presented in Table 3. To assist in their interpretation, it should be noted that an *r* of between 0.10 and 0.29 indicates a weak relationship; 0.30

Table 2: Item test-retest reliabilities for Parts A and B of the ADOM

Measure	n	Test	Retest	*k*	Mean difference (95% CI)	ICC (95% CI)
Part A						
Alcohol						
Mean days used ± *SD*	54	9.0 ± 10.4	7.7 ± 9.3		1.3 (-0.1, 2.7)	0.87 (0.79, 0.92)
Mean units on a typical day ± *SD*	54	6.9 ± 8.6	6.1 ± 7.1		0.8 (-0.4, 2.0)	0.84 (0.74, 0.90)
Primary substance: *n* (%)	44	22 (50)	22 (50)	0.92		
Cannabis						
Mean days used ± *SD*	55	8.9 ± 12.1	8.1 ±11.5		0.9 (-0.3, 2.0)	0.93 (0.89, 0.96)
Primary substance: *n* (%)	44	4 (9.1)	5 (11.4)	0.78		
Amphetamine-type stimulants						
Mean days used ± *SD*	55	1.0 ± 3.1	0.8 ± 2.6		0.3 (0.1, 0.5)	0.96 (0.93, 0.98)
Primary substance: *n* (%)	44	4 (9.1)	4 (9.1)	0.88		
Opiods						
Mean days used ± *SD*	55	7.4 ± 12.0	7.8 ± 12.1		-0.5 (-1.5, 0.6)	0.95 (0.91, 0.97)
Primary substance: *n* (%)	44	17.(38.6)	17 (38.6)	1.00		
Sedatives/ tranquilisers						
Mean days used ± *SD*	53	1.3 ± 4.7	0.8 ± 3.9		0.5 (-0.1, 1.2)	0.86 (0.77, 0.92)
Primary substance: *n* (%)	44	1 (2.3)	1 (2.3)	1.00		
Cigarettes/ Nicotine						
Mean units on a typical day ± *SD*	53	13.7 ± 12.2	13.0 ± 12.7		0.7 (-.04, 1.8)	0.95 (0.91, 0.97)
Primary substance: *n* (%)	44	4 (9.1)	6 (13.6)	0.49		
Injected drug use						
Mean days used ± *SD*	49	7.0 ± 11.7	6.8 ± 11.3		0.2 (-0.4, 0.9)	0.98 (0.97, 0.99)
Shared equipment: *n* (%)	53	1 (1.9)	2 (3.8)	0.49		
Part B						
Mean 'physical health' rating (0-4) ± *SD*	54	1.6 v 1.5	1.2 ± 1.3	0.52		
Mean 'psychological health' rating (0-4) ± *SD*	52	2.2 ± 1.5	1.7 ± 1.4	0.50		
Mean 'family/ friend conflict' rating (0-4) ± *SD*	55	1.3 ± 1.4	1.1 ± 1.1	0.45		
Mean 'work/ activity interference' rating (0-4) ± *SD*	53	1.5 ± 1.6	1.3 ± 1.3	0.46		
Mean 'engaged in work/ other activity' rating (0-4) ± *SD*	54	2.8 ± 1.6	2.7 ± 1.5	0.60		
Mean 'housing difficulties' rating (0-4) ± *SD*	55	0.4 ± 0.9	0.4 ± 1.1	0.48		
Mean 'illegal activity' rating (0-4) ± *SD*	55	0.7 ± 1.3	0.7 ± 1.3	0.59		

and 0.49, a moderate relationship; and an *r* between 0.50 and 1.0, a strong relationship (Cohen, 1988). As can be seen, all correlations are indicative of a strong relationship.

The concurrent validity of the 'sharing injecting equipment' question on ADOM Part A could not be calculated due to the low response rate on this item.

ADOM Part B

Correlations between comparable parts of ADOM Part B, Sections Three and Four of the TOP, and Questions Four and Five of the SF-36 were calculated, based on data obtained at the baseline interview. Findings are presented in Table 4. As can be seen, the correlation between ADOM Part B measures and the comparable TOP and SF-36 questions were variable. Nevertheless, a large number of strong and moderate correlations were reported and the weaker correlations generally resulted when there was less correspondence between the item contents (due to focus or specificity). It is worth noting that the particularly low correlations between ADOM Part B, question 16 and the TOP question "days attended college or school" (r = 0.08) and ADOM Part B, question 17 and the

Table 3: Correlations (Spearman's r) between comparable parts of ADOM Part A, the DDI and the TLFB

Measure	Comparison Instrument			
	TLFB		DDI‡	
	r	*p*	*r*	*p*
Mean days used				
Alcohol	0.96	<0.001	-	-
Cannabis	0.93	<0.001	0.97	<0.001
Amphetamine-type stimulants	0.93	<0.001	-	-
Opioids	0.98	<0.001	-	-
Sedatives/ tranquilisers	0.90	<0.001	0.77	<0.001
Injected drug use	-	-	0.98	<0.001
Mean units on a typical day				
Alcohol	0.94	<0.001	0.89	<0.001
Cigarettes/ Nicotine	-	-	0.88	<0.001

‡Concurrent validity between ADOM and DDI 'mean days used' alcohol, opioids or amphetamine-type stimulants measures could not be calculated due to the scoring formats of the respective instruments.

TOP question "at risk of eviction" (r = 0.16) were most likely due to their low rate of endorsement; only three and five participants endorsed these options, respectively, on the baseline TOP. The concurrent validity of ADOM Part B, question 14 was not measured due to the lack of a suitable comparison question on either the TOP or SF-36.

Sensitivity to Change

All baseline measures were re-administered to participants after a mean of 33.9 days (SD = 10.7; range 18 – 86 days).

ADOM Part A

The results of a series of paired samples t-tests comparing baseline scores with follow-up scores on a range of ADOM Part A questions are presented in Table 5. Table 5 also presents the

Table 4. Correlations (Spearman's r) between comparable parts of ADOM Part B, the TOP and the SF-36

Item	*r*	*p*
ADOM Q12 Physical health		
TOP, Section 4, d - Clients rating of physical health status	-0.36	<0.01
SF-36, Q4, a - Cut down on amount of time spent on work/other activities due to physical health	0.53	<0.001
SF-36, Q4, b - Accomplished less than you would like due to physical health	0.38	<0.01
SF-36, Q4, c - Were limited in the kind of work/other activities due to physical health	0.56	<0.001
SF-36, Q4, d - Had difficulty performing work/other activities due to physical health	0.53	<0.001
ADOM Q13 Psychological health		
TOP, Section 4, a - Clients rating of psychological health status	-0.53	<0.001
SF-36, Q5, a - Cut down on amount of time spent on work/other activities due to emotional problems	0.52	<0.001
SF-36, Q5, b - Accomplished less than you would like due to emotional problems	0.54	<0.001
SF-36, Q5, c - Didn't do work/other activities as carefully as usual due to emotional problems	0.55	<0.001
ADOM Q15 Work/activity interference		
SF-36, Q4, a - Cut down on amount of time spent on work/other activities due to physical health	0.41	<0.01
SF-36, Q4, b - Accomplished less than you would like due to physical health	0.24	0.06
SF-36, Q4, c - Were limited in the kind of work/other activities due to physical health	0.38	<0.01
SF-36, Q4, d - Had difficulty performing work/other activities due to physical health	0.4	<0.01
ADOM Q16 Engaged in work/other activity		
TOP, Section 4, b - Days paid work	0.56	<0.001
TOP, Section 4, c - Days attended college or school	0.08	0.55
ADOM Q17 Housing difficulties		
TOP, Section 4, e - Acute housing problem	0.6	<0.001
TOP, Section 4, f - At risk of eviction	0.16	0.21
ADOM Q18 Illegal activity		
TOP, Section 3, a - Shoplifting	0.18	0.16
TOP, Section 3, b - Drug selling	0.41	<0.01
TOP, Section 3, d - Property theft or burglary	0.21	0.1
TOP, Section 3, f - Committing assault or violence	0.18	0.17

Table 5. Sensitivity to change for ADOM Part A scaled items and comparative TLFB scaled items

Measure	ADOM Part A				TLFB			
	Interview		*p*	RCI%	Interview		*p*	RCI%
	Test	Follow-Up			Test	Follow-Up		
Mean days used ± *SD*								
Alcohol	8.2 ± 9.9	5.8 ± 8.0	0.01	41.8	7.3 ± 9.3	6.5 ± 8.6	0.38	42.9
Cannabis	9.0 ± 11.9	7.6 ± 11.4	0.11	32.1	8.5 ± 11.4	6.9 ± 11.0	0.07	32.1
Amphetamine-type stimulants	0.8 ± 2.3	0.6 ± 1.8	0.12	16.4	0.7 ± 2.0	0.4 ± 1.5	0.2	19.6
Opioids	9.0 ± 12.7	2.5 ± 6.6	0.01	28.6	9.5 ± 12.8	2.7 ± 7.2	0.01	32.1
Sedatives/tranquilisers	1.4 ± 4.7	0.4 ± 1.5	0.05	14.3	1.3 ± 4.2	0.3 ± 1.7	0.01	14.3
Injected drug use	9.1 ± 12.7	2.6 ± 6.9	0.01	30.0				
Mean units on typical using day ± *SD*								
Alcohol	5.9 ± 6.8	3.9 ± 4.5	0.04	49.1	5.6 ± 7.1	4.4 ± 5.1	0.23	50
Cigarettes/Nicotine	14.9 ± 11.8	13.5 ± 11.8	0.09	38.9	10.5 ± 12.2	10.4 ± 12.5	0.97	37.8

percentage of participants whose reported change (between baseline and follow-up) was greater than the Reliable Change Index (RCI) at $p<0.05$.

As can be seen, statistically significant reductions in mean days of use were reported for alcohol, opioids, sedatives/tranquilisers, and injected drug use. A statistically significant reduction was also reported in the mean number of standard drinks consumed per drinking day. The percentage of participants whose change was greater than the RCI ranged from a low of 14.3 (mean days use sedatives/tranquilisers) to a high of 49.1 (mean number of standard drinks consumed per drinking day).

Table 5 also presents data on a range

Table 6. Sensitivity to change for ADOM Part B items

Question	Participant Response					*p*	% Change†
	never	< weekly	1-2 x week	3-4 x week	Daily		
Q12. Physical Health							
Baseline: *n* (%)	17 (30.9)	13 (23.6)	9 (16.4)	5 (9.1)	11 (20.0)		
Follow-Up: *n* (%)	22 (39.3)	15 (26.8)	13 (23.2)	5 (8.9)	1 (1.8)	0.02	63.6
Q13. Psychological Health							
Baseline: *n* (%)	10 (18.2)	12 (21.8)	8 (14.5)	7 (12.7)	18 (32.7)		
Follow-Up: *n* (%)	20 (35.7)	15 (26.8)	12 (21.4)	1 (1.8)	8 (14.3)	0.01	81.8
Q14. Family/Friend Conflict							
Baseline: *n* (%)	22 (39.3)	16 (28.6)	7 (12.5)	3 (5.4)	8 (14.3)		
Follow-Up: *n* (%)	36 (64.3)	13 (23.2)	5 (8.9)	0 (0)	2 (3.6)	0.18	57.1
Q15. Work/Activity Interference							
Baseline: *n* (%)	25 (45.5)	4 (7.3)	8 (14.5)	6 (10.9)	12 (21.8)		
Follow-Up: *n* (%)	33 (60.0)	10 (18.2)	8 (14.5)	2 (3.6)	2 (3.6)	0.16	49.1
Q16. Engaged in Work/Other Activity							
Baseline: *n* (%)	11 (19.6)	3 (5.4)	4 (7.1)	8 (14.3)	30 (53.6)		
Follow-Up: *n* (%)	10 (17.9)	1 (1.8)	12 (21.4)	6 (10.7)	27 (48.2)	0.5	37.5
Q17. Housing Difficulties							
Baseline: *n* (%)	48 (85.7)	3 (5.4)	2 (3.6)	1 (1.8)	2 (3.6)		
Follow-Up: *n* (%)	49 (87.5)	3 (5.4)	1 (1.8)	1 (1.8)	2 (3.6)	0.94	17.9
Q18. Illegal Activity							
Baseline: *n* (%)	38 (67.9)	5 (8.9)	6 (10.7)	2 (3.6)	5 (8.9)		
Follow-Up: *n* (%)	47 (83.9)	3 (5.4)	5 (8.9)	1 (1.8)	0 (0)	0.47	32.1

† Defined as the percentage of participants whose response changed by one or more categories (e.g. from 'never' to '<weekly') between baseline and follow-up (inclusive of both positive and negative change).

of TLFB questions for comparative purposes. As can be seen, statistically significant reductions in mean days of use were reported for opioids and sedatives/tranquilisers. This was consistent with the ADOM Part A data; however, unlike the ADOM Part A data, statistically significant reductions were not identified for mean days of alcohol use or mean number of standard drinks consumed per drinking day. The percentage of participants whose change was greater than the RCI was highly consistent between the measures on all items, including the two alcohol-related items, suggesting this may be the better indicator of sensitivity to change.

ADOM Part B

The results of a series of Wilcoxin Signed Ranks tests comparing baseline scores with follow-up scores on ADOM Part B questions are presented in Table 6. Table 6 also presents the percentage of participants whose response changed by one or more categories between baseline and follow-up (% Change). As can be seen, statistically significant changes were reported for the 'physical health' and 'psychological health' questions. The percentage of participants whose response changed by one or more categories between baseline and follow-up ranged from a low of 17.9 for the 'housing difficulties' question to a high of 81.8 for the 'psychological health' question.

Table 7 presents the results of a series of paired samples t-tests comparing baseline scores with follow-up scores on a range of TOP questions (continuous data) for comparative purposes . The reported improvement in psychological and physical health is consistent with the ADOM Part B data. The reported improvement in overall quality of life is not directly comparable with any ADOM Part B questions, but is suggestive of improvement in multiple life areas. No significant differences in respect to reported mean days of shoplifting, drug selling, paid work, or school attendance, is consistent with ADOM Part B data.

As a means of further comparison, McNemar tests were conducted to compare possible changes (between baseline and follow-up) in participant SF-36 responses. Statistically significant changes were identified on three of the four questions pertaining to physical health: cut down the amount of time spent on work or other activities (p = 0.03), accomplished less than you would have liked (p =0.001), had difficulty performing work or other activities (p = 0.03). Similarly, statistically significant changes were identified on two of the three questions pertaining to psychological health: accomplished less than you would like (p = 0.04), didn't do work or other activities as carefully as usual (p = 0.008). In all cases, the reported change was generally in a positive direction. These findings are consistent with the ADOM Part B.

Table 7. Sensitivity to change for comparative TOP scale items

Measure	Interview		*p*
	Test	Follow-Up	
Mean days (0-28)			
Shoplifting	0.02 ± 0.14	0.04 ± 0.27	0.32
Drug selling	1.1 ± 4.4	1.1 ± 4.5	1
In paid work	7.8 ± 9.3	7.2 ± 9.5	0.56
Attending school	0.8 ± 3.3	0.6 ± 2.8	0.42
Mean rating (0-20)			
Psychological health	11.0 ± 5.5	12.6 ± 5.1	0.02
Physical health	11.5 ± 4.7	13.8 ± 4.3	0.01
Overall quality of life	10.9 ± 5.0	14.6 ± 3.9	0.01

DISCUSSION

The aim of this study was to develop a brief outcome monitoring instrument suited for routine use with clients in the New Zealand AOD treatment sector. It has resulted in the creation of the Alcohol and Drug Outcome Measure (ADOM), an 18-item two part questionnaire. Analyses of the test-retest reliability, concurrent validity and sensitivity to change of Part A of the ADOM consistently produced excellent results and the comparable test results for Part B of the instrument were generally above minimum acceptable standards. The less impressive performance of Part B of the ADOM was not unexpected. Cohen's Kappa was a more conservative measure of test-retest reliability in the context of the Part B questions, as compared to the Part A questions, due to the greater number of response possibilities (thereby providing more opportunity for disagreement). Tests of concurrent validity and sensitivity to change were also hampered due to the imperfect match between a number of the ADOM Part B questions and the respective comparison measures. Matching of the Part A questions was consistently better, due in large part to the greater objectivity of the Part A questions. The poorer performance of Part B in part also reflects a universal challenge faced by such an instrument, the conflict between brevity and scope. Part A consists of a series of items relating to the domain of substance use, with 11 fairly specific questions, while Part B attempts to capture a wide range of other domains in only seven questions, many of which are necessarily substantially broader in scope.

The ADOM development process was highly successful in many respects: the questionnaire content was informed by extensive consultation with a wide range of AOD treatment stakeholders; clients of Māori and Pacific Island ethnicity were well represented in the statistical analyses; testing was conducted in 'real world' treatment services with a mix of client groups under 'real world' conditions; and there was minimal dropout in the psychometric testing sample. These factors together with the satisfactory psychometric findings suggest that with further refining the ADOM has the potential to demonstrate differences in client severity across groups and differential outcomes across treatment approaches and settings.

Nevertheless, a number of limitations are acknowledged. These include: the psychometric testing sample was not obtained by random selection;

not all AOD treatment modalities were represented in the psychometric testing sample (e.g. inpatient detoxification); comparison measures used for Part B of the ADOM were not always well matched; and sensitivity to change was measured after a relatively brief period post-admission. These limitations were largely the result of budget and time constraints, as well as the trade-offs that frequently occur when conducting research in a clinical environment, although the short time-frame did have the advantage of helping achieve the high follow-up rate.

Despite these limitations, it is reasonable to conclude that ADOM Part A has excellent potential as a measure of type and frequency of AOD use. Confidence in Part B could be increased if those questions with less impressive statistics were removed from the instrument. Candidate questions for removal, due to their especially low concurrent validity and relatively low sensitivity to change, might include: 'how often has your alcohol or drug use led to conflict with friend or family members', 'how often have you had difficulties with housing or finding somewhere stable to live', and 'how often have you been involved in any criminal or illegal activity'. However, poor test results in these areas were most likely the result of imperfectly matched comparison measures (as discussed above) and/or, in the case of the housing difficulty and illegal activity questions, irrelevance to the majority of clients who participated in this trial.

It is also worth highlighting that whilst an improvement on key outcome measures is frequently evidenced early in treatment for many clients, for others stabilisation is realistic treatment progress and for others improvement may happen more slowly. Thus, progress should not always be expected in the early stages of treatment. Outcome monitoring is of importance in this context as, if no change is identified, or in the case of negative progress the treatment plan can be reviewed and amended as required (Teesson et al., 2000). All of these considerations suggest that it is premature and/or ill-considered to remove questions from ADOM Part B at this time.

In light of these considerations, thought could be given to employing a refined version of the ADOM as a core outcome monitoring instrument in the AOD sector with additional brief modules for particular client groups. For example, injecting drug use, stability of housing and criminal activity are highly relevant outcome measures for clients who present for treatment with dependence on illegal substances and are typically included in brief outcome measures targeted for use with opioid substitution treatment clients and other illicit drug user groups (Deering et al., 2008; Marsden et al., 2008).

In reflecting on the potential utility of ADOM Part B, it is also worth considering the practical realities of routine outcome monitoring in a clinical environment. If the aim is to produce a single outcome monitoring instrument for use across all AOD treatment modalities then a number of "tradeoffs" need to be carefully considered in regard to implementing and sustaining the use of a brief, generic client outcome measure across a range of client groups in real life clinical settings. Obtaining "buy in" from clinicians and clients is of critical importance which therefore places a high priority on feasibility characteristics (Slade, Thornicroft, & Glover, 1999). Typically, instruments that score high in regard to psychometric properties of reliability and validity are low on clinical feasibility and vice versa. On the one hand, brief questionnaires administered as a structured interview with the flexibility to enable more in depth exploration or digression as required, may not be as psychometrically sound as longer instruments. On the other hand, they may have good clinical utility and therefore provide a degree of measurement precision that is realistic and clinically useful and which can lead to improvements in the quality of care (Berwick et al., 1991). Taking into account that developing a clinically useful instrument was prioritised the outcome of the psychometric testing process was largely consistent with these considerations.

Furthermore while the need for routine outcome monitoring is evident and strongly advocated, the actual implementation of outcomes monitoring within AOD clinical settings needs to be addressed. While it is necessary as a first step to develop valid and reliable outcome measures that are considered suitable for use by clients and clinicians and to identify how to improve treatment outcomes in theory, these strategies need to actually work in practice. This critical need to investigate the implementation of routine outcome measures within real life clinical settings, particularly their perceived value and utility, is borne out by overseas research findings. For example, staff attitudes and experiences of outcomes monitoring in every day clinical practice together with lack of attention to implementation have been shown to influence practice and completion rates (Trauer, Callaly, & Herrman 2009).

In conclusion, taking all of the above considerations into account, it is the recommendation of the project team that the ADOM is field tested in a small number of AOD treatment services. The level of uptake and perceived clinical utility should be closely examined and possible refinements of Part B identified. To maximise its clinical utility, current and previous ADOM results should be readily accessible to clients and clinicians involved in the field testing process. A sound training program, instructive resources, supportive data collection and management systems, and the provision of ongoing support should also be central platforms of any attempt to test the ADOM in the AOD treatment sector. As commented above, failure to address implementation is highly likely to undermine the potential of the ADOM to serve as a potentially beneficial core outcomes monitoring instrument. Conversely, careful attention to the implementation phase provides the opportunity for clinicians and clients to trial Part A and Part B of the ADOM and to supplement one or other, or both, with other measures relevant to the service or therapeutic context.

Acknowledgements

This study was made possible through funding from Te Pou, the National Centre of Mental Health Research, Information and Workforce Development. The authors acknowledge the members of the project advisory board and the staff and clients of those services who so willingly participated in the study.

References

Berwick, D. M., Murphy, J. M., Goldman, P. A., Ware Jnr, J. E., Barsky, A. J., & Weinstein, M. C. (1991). Performance of a five-item mental health screening test. *Medical Care, 29*(2), 169-176.

Christie, G., Marsh, R., Sheridan, J., Wheeler, A., Suaalii-Sauni, T., Black, S., et al. (2007). The Substances and Choices Scale (SACS) - the development and testing of a new alcohol and other drug screening and outcome measurement instrument for young people. *Addiction, 102*(9), 1390-1398.

Cohen, J. W. (1988). *Statistical power analysis for the behavioural sciences* (2nd ed.). Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Deering, D. E., Robinson, G., Adamson, S., Paton-Simpson, G., Robertson, P., Warren, H., et al. (2004). *Alcohol and Drug Outcomes Project (ADOPT)*. Auckland: Health Research Council of New Zealand. Retrieved from http://www.tepou.co.nz/file/PDF/publications/3_63_119_ADOPT%20Final%20Report%20August%2004.pdf

Deering, D. E., Sellman, J. D., Adamson, S. J., Horn, J., & Frampton, C. M. A. (2008). Development of a brief treatment instrument for routine clinical use with methadone maintenance treatment clients: The Methadone Treatment Index. *Substance Use and Misuse, 43*(11), 1547-1561.

Fleiss, J. L. (1991). *Statistical methods for rates and proportions.* New York: John Wiley.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data models. *Biometrics, 33*(1), 159-174.

Lawrinson, P., Copeland, J., & Indig, D. (2005). Development and validation of a brief instrument for routine outcome monitoring in opioid maintenance pharmacotherapy services: the Brief Treatment Outcome Measure (BTOM). *Drug and Alcohol Dependence, 80*(1), 125-133.

Lawrinson, P., Roche, A., & Copeland, J. (2009). Optimizing clinical care through implementation of outcome monitoring systems. *Drugs: Education, Prevention and Policy, 16*(3), 250-259.

Marsden, J., Farrell, M., Bradbury, C., Dale-Perera, A., Eastwood, B., Roxburgh, M., et al. (2008). Development of the treatment outcomes profile. *Addiction, 103*(9), 1450-1460.

McLellan, A., Mckay, J., Forman, R., Cacciola, J., & Kemp, J. (2005). Reconsidering the evaluation of addiction treatment: From monitoring retrospective follow-up to concurrent recovery monitoring. *Addiction, 100*(4), 447-458.

Medical Outcomes Trust. (1994). *Australia/New Zealand SF-36 health survey*. Medical Outcomes Trust: The Health Institute, Boston.

Simpson, M., Lawrinson, P., Copeland, J., & Gates, P. (2009). The Alcohol Treatment Outcome Measure (ATOM): a new clinical tool for standardising outcome measurement for alcohol treatment. *Addictive Behaviors, 43*(1), 121-124.

Slade, M., Thornicroft, G., & Glover, G. (1999). The feasibility of routine outcome measures in mental health. *Social Psychiatry and Psychiatric Epidemiology, 34*(5), 234-249.

Sobell, L. C., & Sobell, M. B. (1996). *Timeline followback user's guide: A calendar method for assessing alcohol and drug use*. Toronto: Addiction Research Foundation.

Teesson, M., Clement, N., Copeland, J., Conroy, A., & Reid, A. (2000). *The measurement of outcome in alcohol and other drug treatment: A review of available instruments*. Kensington, NSW: National Drug and Alcohol Research Centre, University of New South Wales.

Trauer, T., Callaly, T., & Herrman, H. (2009). Attitudes of mental health staff to routine outcome measurement. *Journal of Mental Health 18*(4), 288-296.

Ware, J. E., & Sherbourne, C. (1992). The MOS 36-Item Short-Form Health Survey (SF-36). *Medical Care, 30*(6), 473-481.

Corresponding Author:

Justin Pulford

Research Fellow,

Gambling and Addictions Research Centre

AUT University

Private Bag 92006

Auckland

New Zealand.

justin.pulford@aut.ac.nz

The Alcohol and Drug Outcome Measure (ADOM): Part A

All questions relate to the past four weeks
The questions do not apply to prescribed medication; however, any misuse of prescription medication should be included e.g. taking more than prescribed/injecting of medications not intended to be injected
If the client has been an inpatient or in custody for more than 22 days during the last four weeks, do not complete the questionnaire.

IN THE PAST FOUR WEEKS:	Days used (0-28)	
1. On how many days did you drink alcohol?		
2. How many standard drinks did you consume on a typical drinking day? (*1 Standard Drink = 1 can of beer, 100ml wine, or1 double of spirits, where bottle of wine = 7 or jug of beer = 3 or 750ml spirits = 23)*		

IN THE PAST FOUR WEEKS, ON HOW MANY DAYS DID YOU USE:

	Days used (0-28)	
3. Cannabis		
4. Amphetamine-type stimulants e.g. methamphetamine, speed, methylphenidate (Rubifen)		
5. Opioids		
6. Sedatives/tranquilisers E.g. diazepam (Valium), temazepam		
7. Any other drugs. e.g. ecstasy, hallucinogens, solvents, GHB etc Specify what drugs: _____ *(interviewer: if "other drugs" contains substances covered in the above questions please return to the appropriate question and recode)*	_____ _____ _____	_ _ _
8. How many cigarettes have you smoked per day, on average (if non-smoker, enter zero):		

9. Please put a tick in the right hand column to identify main substance of concern (for some clients there may be more than one).

IN THE PAST FOUR WEEKS:

	Days injected (0-28)	
10. On how many days have you injected drugs? *(if none, enter zero and go to question 12)*		
11. Have you shared any injecting equipment? (*sharing means using someone else's equipment which has already been used or someone using yours regardless of whether you were both present at the time or not; equipment includes needles, syringes, water, dregs, tourniquets, spoons, filters)*	Yes	No

Please turn over

The Alcohol and Drug Outcome Measure (ADOM): Part B

IN THE PAST FOUR WEEKS:

1. How often has your physical health interfered with your day-to-day functioning?

Never	Less than weekly	Once or twice a week	Three or four times a week	Daily or almost daily

2. How often has your psychological or mental health interfered with your day-to-day functioning?

Never	Less than weekly	Once or twice a week	Three or four times a week	Daily or almost daily

3. How often has your alcohol or drug use led to conflict with friends or family members?

Never	Less than weekly	Once or twice a week	Three or four times a week	Daily or almost daily

4. How often has your alcohol or drug use interfered with your work or other activities (include social, recreational, parenting/caregiving, study or other personal activities)?

Never	Less than weekly	Once or twice a week	Three or four times a week	Daily or almost daily

5. How often have you engaged in paid employment, voluntary work, study, parenting or other care giving activities?

Never	Less than weekly	Once or twice a week	Three or four times a week	Daily or almost daily

6. How often have you had difficulties with housing or finding somewhere stable to live?

Never	Less than weekly	Once or twice a week	Three or four times a week	Daily or almost daily

7. Apart from using illicit substances, how often have you been involved in any criminal or illegal activity (e.g. driving a motor vehicle under the influence of alcohol or drugs or supplying an illicit substance to another person)?

Never	Less than weekly	Once or twice a week	Three or four times a week	Daily or almost daily