International Journal of Mental Health and Addiction
https://doi.org/10.1007/s11469-021-00630-y

ORIGINAL ARTICLE

Cultural Adaptation and Psychometric Validation of the Australian Treatment Outcomes Profile for the Greek Population

Paraskevi Karakoula[1] · Konstantinos Kokkolis[1] · Charikleia Tsatsaroni[1] · Aggeliki Kitsiou[1] · Emmanouil Koukidis[1] · Paraskevi Danousi[1] · Sofia Nikolaidi[1] · Georgios Kolios[1] · Georgia Vlachou[1] · Apostolia Fragkouli[1] · Nicholas Lintzeris[2,3,4]

Accepted: 20 August 2021

Abstract
The study aimed to culturally adapt the Australian Treatment Outcomes Profile (ATOP; Ryan et al., 2014) in the Greek context and to evaluate its psychometric properties in a sample of people in substitution treatment for opiate dependence. The Hellenic Treatment Outcomes Profile (HTOP) was initially validated through the process of clinician and patient focus groups that ensured its content and face validity, respectively. Inter-rater reliability and concurrent validity were satisfactory for all HTOP items. The HTOP is an extremely useful instrument that assesses the patients' therapeutic course briefly and comprehensively. We expect it will contribute to a more targeted and individualized patient care, as well as to a better evaluation of our services, through standardizing the therapeutic outcomes monitoring.

Keywords Hellenic Treatment Outcomes Profile · Opioid substitution treatment · Cultural adaptation · Psychometric validation

Substance abuse is a major public health issue in every developed and developing country. It has been a major societal and health concern in Greece for decades, especially during the economic crisis of the last decade, which gave rise to increased substance use (Thomaidis et al., 2016) and, consequently, the spread of the human immunodeficiency virus (HIV/AIDS) and the hepatitis C virus (HCV) (Paraskevis et al., 2013). The Organization Against Drugs (OKANA), the public provider for opioid substitution therapy in Greece operating

✉ Paraskevi Karakoula
pkarakoula@okana.gr; evi_karakoula@yahoo.gr

[1] Organization Against Drugs (OKANA), Averof 21, 104 33 Athens, Greece

[2] Sydney School of Medicine (Central Clinical School), Faculty of Medicine and Health, The University of Sydney, Sydney, NSW, Australia

[3] Drug and Alcohol Services, South Eastern Sydney Local Health District, Sydney, NSW, Australia

[4] NSW Drug and Alcohol Clinical Research and Improvement Network (DACRIN), Sydney, NSW, Australia

since 1995, extended its service network during the crisis, in order to address the rise in both substance use and related infectious diseases (HIV/HCV), as well as to provide health care for the opioid dependent population, a vulnerable group significantly affected by the crisis. Indeed, this expansion contributed in containing the AIDS epidemic in 2011–2013 (Malliori et al., 2013) and absorbing waiting lists for substitution therapy until 2017, thus achieving what had been a previously unreached goal of many years. Although budget and personnel cuts during the crisis did not allow for a systematic evaluation of the services provided, with the economic crisis ending OKANA has set the evaluation of therapeutic outcomes as a new goal for the treatment system.

Therapeutic outcomes monitoring (a) enables the evaluation of our interventions' effectiveness (Affholter, 1994), (b) offers invaluable feedback to the therapists and management resulting in an ongoing quality improvement of the services provided (Hunt et al., 2017), and (c) may prove a useful research tool (Lawrinson et al., 2009; Teruya et al., 2006). To that end, an instrument that can briefly assess substance abuse and general wellbeing of the patients is deemed necessary.

To our knowledge, none of the brief instruments, which have been developed recently to monitor outcomes, have yet been adapted for the Greek population. Moreover, most of the ones available are either too lengthy, for example, the Brief Treatment Outcome Measure (BTOM; Lawrinson et al., 2003; Lawrinson et al., 2005), or not comprehensive enough, like the Brief Addiction Monitor (BAM; Cacciola et al., 2013). Therefore, in order to enhance regular patient assessment, and not just at intake, a brief as well as comprehensive instrument is needed.

The Treatment Outcomes Profile interview (TOP; Marsden et al., 2008) was developed in the UK in order to monitor the clinical outcome of drug and alcohol services. The TOP briefly assesses recent substance use behavior, injecting behavior, social functioning (defined by work, education, housing), criminal activity, and overall wellbeing (psychological and physical health, quality of life). The TOP has been found to be valid, reliable, and sensitive to change (Marsden et al., 2008). It has already been adapted for the Australian (ATOP; Ryan et al., 2014), Chilean (Castillo-Carniglia et al., 2015), and Chinese (Wang et al., 2017) population. The ATOP has been also validated for use in patients with cannabis dependence (Mills et al., 2020) and for telephone administration in drug health treatment populations (Deacon et al., 2020).

The Australian adaptation of the TOP (ATOP) demonstrates a slightly different structure. It is comprised of two sections that assess recent substance use and injecting behavior, as well as health and well-being. It also has shown very good validity, reliability, and clinician ratings (Ryan et al., 2014). We chose to adapt this version of the instrument in the Greek context for many reasons. First of all, the ATOP does not include Sect. 3 that assesses criminal activity, due to its low inter-rater reliability in the Australian version (Ryan et al., 2014) and the questionable validity of its criminal activity items (Luty et al., 2009). According to the traditions of Greek Addiction Treatment Programs, criminal activity is not considered an essential aspect of the therapeutic procedures, and the recent national legislations are congruent to this mentality, in order to minimize the addiction associated stigma (Law 4139/2013, as amended and in force); therefore, excluding criminal activity assessment seemed more suitable to the Greek context. Second, the substance use profile of the Australian version is closer to the Greek reality, especially concerning substance users' preferences for opioids and benzodiazepines, according to recent drug use prevalence estimations in the Greek population (EPIPSI & EKTEPN) 2018). Third, Sects. 1 (substance use) and 2 (injecting risk behavior) were condensed in one section (substance use and injecting risk behavior), and the self-rating health and quality of life

scales were gathered in a sub-section of Sect. 2 (health and wellbeing), further shortening the interview. Lastly, the self-rating scales of the ATOP are 10-point, as opposed to the 20-point self-rating scales of the TOP, which we concur is the health rating scale the Greek population is most accustomed to.

The study aimed to culturally adapt the Australian Treatment Outcomes Profile (ATOP; Ryan et al., 2014) in the Greek context and to evaluate its psychometric properties in a sample of people in substitution treatment for opiate dependence. In particular, we tested its inter-rater reliability, concurrent validity, and change sensitivity.

Material and Methods

Adaptation of the Instrument to the Greek Context

The adaptation process began with the translation of the ATOP, which was undertaken by a panel of experts on substance abuse with a research background. The panel was comprised of one psychiatrist, two psychologists, two sociologists, and a nurse. All but one worked on the forward translation and, after unanimously consenting on it, the member that had not participated so far proceeded with the backward translation. The original questionnaire was compared with the backward translation, and differences were smoothed out with a focus on conceptual and cultural appropriateness, as opposed to linguistic accuracy. This version was reviewed by the lead author of the ATOP, Prof. Lintzeris, and then verified through the process of focus groups.

Five focus groups were conducted next, two comprised of therapists working at various substance dependence rehabilitation programs of OKANA and three comprised of patients attending various programs and representing various age groups. Each one of the therapists' focus groups was comprised of two psychiatrists, two psychologists, two social workers, and two nurses and was coordinated by two members of the research team. Therapists were employed in substitution programs, adolescents' program, substitution and cannabis withdrawal program, and social reintegration program, and sexes were represented equally. Our aim was to ensure the instrument's clinical utility and content validity through their feedback on the interview. Two of the patients' focus groups were attended by seven and eight patients, respectively, polydrug users, with heroin being their drug of choice, aged 30–55 years old, in order to represent our "typical" patients. There was one woman in each group, as women tend to be underrepresented in our programs. In the third patients' focus group, we included seven adolescents and young adults, problematic users of various substances who had not been using substances for more than 10 years and were not high-risk drug users (i.e., intravenous drug users). Patients' focus groups significantly assisted in ensuring the interview's clarity and face validity. Finally, a pilot administration in 10 volunteer patients was conducted, in order to check the administration protocol, instructions, time needed, etc. Members of the research team administered the entire battery of instruments to each volunteer individually and concluded that instructions were clear and average administration time was 12' for the HTOP and 45' for the entire protocol.

Certain additions and modifications emerged from these procedures. First, an item assessing sports and/or volunteer work was added to work and education as an indicator of social functioning, resulting from the adolescents' focus group and their input on their age group's range of social activities. Second, we added cells to record the route of administration for each substance used, due to the varying degree of harm and clinical importance of

different administration routes, but also in order to epidemiologically record the observed increase in methamphetamine use (sisa) with improvised shared inhalers, which could lead to a new increase of infectious disease (HIV, HCV), as these inhalers cause lacerations around the mouth (Fischer et al., 2008). Third, we changed the sequence of the two questions that regard violence, after systematically observing that some participants who hesitated at first to admit being violent themselves were more open about their violent behaviors after having answered about being on the receiving end of such behaviors. Changes are depicted in Table 1.

Therefore, after evaluating content validity of the HTOP through expert opinions, and its face validity through participants' judgments, we concluded that the Greek version of the instrument (Hellenic Treatment Outcomes Profile (HTOP)) was ready to be administered in the patients using substance dependence rehabilitation services in Greece.

Procedures and Psychometric Evaluation

Prior to proceeding with the study, our research team conducted simulation sessions, in order to standardize the administration and scoring procedures. Each member of the research team administered the HTOP either to another team member or a volunteer colleague, in the presence of the entire team. During every simulation interview, the rest of the team was scoring the interviewee's answers, and after the interview, we discussed issues pertaining to the administration procedure and compared scores in order to ensure we all abided to the same scoring rules and reasoning.

The main study included three stages: an initial test and a re-test 2–7 days later. In the initial test stage, participants were also administered four other questionnaires to test the concurrent validity of HTOP items: the Alcohol Use Disorders Identification Test (AUDIT-c; Bush et al., 1998) for the number of days of alcohol use in the last 28 days; the Greek version of the EuropASI (EuropASI Working Group, 1994; Kokkevi & Hartgers, 1995; EPIPSI & EKTEPN, 1996) for the number of days of injecting drug use in the last 28 days; the Greek version of the Medical Outcome Study (MOS) 36-Item Short-Form Health Survey (SF-36; Anagnostopoulos et al., 2005; Pappa et al., 2005; Ware & Sherbourne, 1992) for physical and psychological health in the last 28 days; and the Greek version of the World Health Organization Quality of Life Instruments (WHO-QOL-BREF; Ginieri-Coccossis et al., 2012; WHOQOL Group, 1998) for the quality of life, as well as psychological and physical health in the last 28 days. Toxicology reports from urine tests were also used to test the criterion validity of the HTOP items referring to substance use. We randomly selected 75% of the participants and used, with their permission, the toxicology tests that were performed as part of their therapeutic protocol the week before the initial test. The initial test battery was administered by a member of the research team who inquired participants about their experiences in the 4 weeks preceding the test. A different researcher conducted the re-test interview with the same reference period for the questions as the initial test.

Participants

We performed a priori power analysis to estimate the required sample size for inter-rater reliability using Kraemer and Thiemann's critical effect size Δ (Kraemer & Thiemann, 1987). With regard to continuous variables, we set Hp at 0.80 (highly correlated

Table 1 Summary of changes made to ATOP

Section	Items	Description of change
Section 1: Substance use	Route of administration	Add in of a column to record the route of administration for each substance used due to the varying degree of harm and clinical importance of different administration routes
Section 1: Substance use	Tobacco use	Tobacco use was not measured using a dummy variable but on a weekly base, in the same way each substance used was recorded
Section 2: Health and wellbeing	Sports and/or volunteer work	An item assessing sports and/or volunteer work was added to work and education as an indicator of social functioning, resulting from the adolescents' focus group and their input on their age group's range of social activities
Section 2: Health and wellbeing	Violent behaviors	Change of the sequence of the two questions that regard violence after systematically observing that some participants who hesitated at first to admit being violent themselves, were more open about their violent behaviors after having answered about being on the receiving end of such behaviors

measures) against Hp0 at 0.50 (no association), resulting in a requirement of 36 participants in order to reject the null-hypothesis (power = 0.9, $\alpha = 0.05$, one-tailed). Regarding dichotomous variables, we set Cohen's kappa at the 0.61 threshold for substantial agreement against 0.21 for the null hypothesis, resulting in a requirement of 37 participants for each comparison (power = 0.9, $\alpha = 0.05$, one-tailed) (Landis & Koch, 1977; Marsden et al., 2008; Wang et al., 2017).

Participants were 147 adult patients attending opioid substitution therapy programs in 14 different units of OKANA in Athens, Greece. Sample selection was random. First, we decided which units would participate in the study. We included the admissions unit, the social reintegration unit, the cannabis withdrawal unit, and the buprenorphine prescription unit — as they all are one of a kind — and then we randomly selected 10 out of the 23 substitution units in Athens. Each selected unit prepared a list of patients who was then assigned a code number. Participants were selected from the coded lists. Both unit and patient selections were made using the sampling command on the Excel Data analysis tab. Exclusion criteria were (a) individuals presenting with acute intoxication or withdrawal symptoms at the time of the administration and (b) patients with severe mental problems, acute psychosis, or other cognitive and communication problems.

Instruments

Hellenic Treatment Outcomes Profile-HTOP

The HTOP is a one-page clinician/researcher administered tool comprised of two sections. The first section assesses substance use and injecting risk behavior in the last four weeks. Section 1 examines days used, out of 28, and route of administration for the following substances: alcohol, cannabis, amphetamines, benzodiazepines, heroin, other opioids, cocaine, other psychoactive substances, and tobacco. The second section assesses health and wellbeing and the items regard work, education, sports/volunteer work, housing, violent behavior, arrests, caring for children, psychological and physical health, and quality of life, all measured using one item each in the past 28 days. Higher scores on the substance use items are indicative of more days of use (0–28), whereas higher scores on the health and wellbeing questions (0–10) and the work, education, and sports/volunteer work items (0–28) depict better self-rated heath outcomes and better social functioning, respectively. The items regarding violent behavior, arrests, and caring for children during the last 28 days were measured using dummy variables (yes/no). The final version of the scale is presented in Supplementary Table 1.

Alcohol Use Disorders Identification Test (AUDIT)

The Alcohol Use Disorders Identification Test (AUDIT) is an instrument for the early diagnosis of alcohol consumption and relevant behaviors and problems, which was developed by the World Health Organization (Saunders & Aasland, 1987). In this study, only the first three items of the AUDIT were used. They assess the quantity and frequency of alcohol consumption, and they are practically equivalent to the AUDIT-c subscale (Bush et al., 1998).

European Addiction Severity Index (EuropASI)

The EuropASI (EuropASI Working Group, 1994; Kokkevi & Hartgers, 1995) is an adaptation of the fifth edition of the Addiction Severity Index (ASI; McLellan et al., 1992). It is a semi-structured interview that assesses seven functioning areas (medical, psychiatric, working, legal, family/social, support systems, substance, and alcohol abuse) in individuals that abuse substances (Leonhard et al., 2000; Rosen et al., 2000). The interview has been validated for the Greek population (EPIPSI & EKTEPN, 1996). In this study, only three items (14, 14a, 14b) referring to risk injecting behavior were used.

World Health Organization Quality of Life Instruments (WHO-QOL-BREF)

The WHOQOL-BREF (WHOQOL Group, 1998) is a 26-item version of the 100-item WHOQOL-100 (WHOQOL Group, 1994a, b, 1995) that assesses quality of life in four areas: physical health, psychological health, social relationships, and environment. Items are scored on a 5-point Likert scale, and higher scores indicate better quality of life. It also assesses the Overall quality of life and General health with a single item for each. The instrument has been validated for the Greek population (Ginieri-Coccossis et al., 2012). For the sample used in this study Cronbach's alphas for the WHOQOL-BREF domains were 0.78 for physical health, 0.75 for psychological health, 0.61 for social relationships, and 0.70 for the environment.

Medical Outcome Study (MOS) 36-Item Short-Form Health Survey (SF-36)

The MOS 36-Item Short-Form Health Survey (SF-36**;** Ware & Sherbourne, 1992) is consisted of eight subscales. Physical functioning, role restrictions due to health problems, pain, and general health comprise the physical health scale, and role restrictions due to emotional problems, energy/fatigue, emotional wellbeing, and social functioning comprise the psychological health scale. All items are scored so that higher scores define more favorable health states. The scale has been validated for the Greek population (Anagnostopoulos et al., 2005; Pappa, et al., 2005). For the sample used in this study Cronbach's alphas ranged from 0.58 (role restrictions due to emotional problems) to 0.93 (physical functioning).

Statistical Analysis

Inter-rater Reliability

Inter-rater reliability was estimated using intra-class correlation coefficients (ICC) for continuous variables and the kappa statistic for categorical variables. ICC and kappa values ≥ 0.75 were considered as showing excellent agreement (Cicchetti and Sparrow, 1981). Test and retest HTOP scores were compared to test inter-rater reliability, whereas retest reliability — usually tested for after at least a 2-week period — was not estimated, since the HTOP does not assess stable characteristics, and any changes in addictive behaviors would fallibly be assessed as lack of retest reliability.

Concurrent Validity

The correlations between the HTOP and the additional instruments were assessed using Spearman's rho correlation coefficient (ρ), which ranges $-1 \leq p \leq +1$. Dancey and Reidy's latest performance (2017) for Pearson's and Spearman's coefficients were employed where values ranging from ± 0.9 to ± 1 are indicative of perfect correlation, whereas values between ± 0.7 to ± 0.9, ± 0.4 to ± 0.6, and ± 0.1 to ± 0.3 are indicative of strong, moderate, and weak correlation, respectively. The self-reporting questionnaires and the toxicology reports were contrasted using Cohen's kappa, together with sensitivity, specificity, and positive and negative agreement (Cicchetti and Feinstein, 1990). For the interpretation of the Cohen's kappa values, the cut off points proposed by Viera and Garrett (2005) are employed. More specifically values between 0.99–0.81, 0.80–0.61, 0.60–0.41, 0.40–0.21, 0.20–0.01 and below zero are indicative of almost perfect, substantial, moderate, fair, slight, and less than chance agreement, respectively.

All analyses were done using IBM® SPSS® Statistics 25.0 (2017).

Results

Patient Demographic Characteristics and Substance Use

The initial test assessment completed by 147 patients, out of those 102 (≈ 70%) were randomly selected to participate at the retest, which was held 2–7 days after the initial test ($M = 3.47$, SD = 1.73 days). Comparisons of socio-demographic and clinical variables and variables related to substance abuse between subjects who participated in retest measurements and those who did not. Slight differences were observed in terms of age and use of benzodiazepines. Findings are presented in Supplementary Table II.

Toxicology reports from urine samples were obtained for 113 (≈ 75%) of the patients through random sampling. No statistically significant difference was observed between patients participated in toxicological examinations and those who did not (Supplementary Table II).

Most of the participants (133, 90.5%) attended buprenorphine substitution therapy, while the rest were on methadone. The majority of the participants were male (123, 83.7%), and their age ranged from 21 to 67 years old ($M = 44.24$, SD = 8.67).

Forty-seven patients (32%) had been in therapy for less than 3 months, 79 (53.7%) were interviewed while attending the main phase of their treatment, and 21 (14.3%) were concluding their therapy (abstinent in the social reintegration program/on low dosage monthly prescription buprenorphine).

Regarding substance use in the last 28 days, 71 (48.3%) had used alcohol, 81 (55.1%) cannabis, 16 (10.9%) amphetamines, 85 (57.8%) benzodiazepines, 65 (44.2%) heroin, 21 (14.3%) other opioids, 36 (24.5%) cocaine, and 142 (96.6%) tobacco.

Inter-rater Reliability

Inter-rater reliability was estimated by comparing test and retest scores for the 102 participants that participated both in the initial test and at the retest. Reliability was found excellent for the prevalence of use, days used, and units per day for most substances.

Exceptions were all measures for other opioids, days of amphetamine use and units of alcohol per day, for which reliability was good, and units of benzodiazepines per day for which it was poor. Agreement was also excellent for most social functioning variables (days of paid work, homelessness, caring for children < 5 and > 5, violent behavior, arrests) but fair for being the victim of violence, and poor for being in danger of eviction, and days of school and sports/volunteering. Finally, inter-rater reliability was good for psychological and physical health and quality of life. Results are shown in Table 2.

Concurrent Validity

Concurrent validity was estimated using the initial sample of 147 participants. Correlations between the HTOP psychological health, physical health and quality of life, and scores of the validation measures ranged from fair to good (Table 3). The correlation between the HTOP alcohol use days and AUDIT-c alcohol use was excellent, as well as the correlation between the HTOP injecting use days and shared paraphernalia with the two corresponding items from the EuropASI (Table 3). The results of the toxicology urine tests for the week prior to the initial test were compared with patient self-reports for the same week. Sample sizes differ for each analysis depending on the number of participants that reported using each substance. Results showed excellent sensitivity and specificity and good to excellent agreement for heroin, cannabis, benzodiazepines, and cocaine. Regarding amphetamines and other opioids, specificity was excellent, whereas sensitivity and agreement were poor (Table 4).

Discussion

Monitoring the outcome of substance abuse treatment interventions is essential in order to evaluate and upgrade our services. To that end, a reliable and comprehensive assessment instrument is necessary. In this study, we evaluated the psychometric properties of the Hellenic Treatment Outcomes Profile in a sample of adult opioid dependent individuals in opioid substitution therapy. Overall, results demonstrate fair to excellent inter-rater reliability, concurrent, criterion, and discriminant validity, consistent with findings of other adaptation studies of the instrument (Castillo-Carniglia et al., 2015; Ryan et al., 2014; Wang et al., 2017), resulting in a culturally appropriate Greek instrument.

The findings of the psychometric evaluation of the Greek version of TOP demonstrated acceptable levels of reliability and validity. Test–retest reliabilities and inter-rater reliabilities were deemed high for most substances and social functioning items except of units of benzodiazepines per day and for being the victim of violence, in danger of eviction and days of school and volunteering. A similar pattern of results was observed in the original scale (Marsden et al., 2008) where the item regarding benzodiazepines use was dropped. Furthermore, the abovementioned social functioning items were mentioned before to have problematic reliability levels (Marsden et al., 2008; Ryan et al., 2014) mainly attributed to small sample size.

The concurrent validity of the scale with commonly used "gold standard" validated instruments was acceptable as in all previous validation attempts of the scale (Castillo-Carniglia et al., 2015; Lintzeris et al., 2016; Marsden et al., 2008; Ryan et al., 2014; Wang et al., 2017). Moreover, there was acceptable concordance, specificity, sensitivity, and

Table 2 Inter-rater reliability of HTOP items ($N = 102$)

HTOP item	Test	Retest	K[a]	Mean diff. (95% CI)	ICC[b] (95% CI[c])
Alcohol					
Used *n* (%)	51 (50%)	47 (46.1%)	.86		
Units used ± SD[d]	1.55 ± 3.44	1.58 ± 3.57		− .03 (− .60, .54)	.66 (.54, .76)
Days used ± SD	3.71 ± 7.77	3.97 ± 7.77		− .27 (− .84,.31)	.93 (.86–.95)
Cannabis					
Used *n* (%)	57 (55.9%)	58 (56.9%)	.94		
Units used (g) ± SD	.87 ± 1.32	.74 ± 1.11		.13 (− .04,.30)	.74 (.64, .82)
Days used ± SD	10.27 ± 12.26	9.98 ± 11.96		.12 (− .32, .56)	.98 (.98, .99)
Amphetamines					
Used *n* (%)	12 (11.8%)	12 (11.8%)	1.0		
Units used (g) ± SD	.11 ± .55	.14 ± .62		− .03 (− .08, .02)	.91 (.87, .94)
Days used ± SD	.72 ± 3.26	1.05 ± 4.51		− .33(− .98, .32)	.65 (.52, .75)
Benzodiazepines					
Used *n* (%)	65 (63.75%)	62(61.4%)	.94		
Units used (tab) ± SD	2.18 ± 4.42	1.32 ± 1.74		.85 (.11, 1.60)	.37 (.19, .53)
Days used ± SD	12.13 ± 12.71	11.79 ± 12.72		.34 (− .66, 1.34)	.92 (.89, .95)
Heroin					
Used *n* (%)	49 (48%)	52 (51%)	.86		
Units used (g) ± SD	.41 ± .88	.40 ± .81		.01 (− .08, .09)	.86 (.81, .91)
Days used ± SD	4.98 ± 8.77	4.27 ± 8.20		.71(− .23, 1.64)	.84 (.78, .89)
Other opioids	18(17.8%)	12(11.8%)	.46		
Used *n* (%)					
Units used (tab) ± SD	.34 ± .83	.27 ± .96		.07 (− .08, .22)	.64 (.51, .74)
Days used ± SD	1.06 ± 3.93	.46 ± 1.77		.59 (− .11, 1.30)	.32 (.13, .48)
Cocaine					
Used n(%)	27 (26.5%)	25 (24.5%)	.85		
Units used (g) ± SD	.35 ± 1.30	.43 ± 1.89		− .08 (− .29, .14)	.77 (.68, .84)
Days used ± SD	1.77 ± 5.66	1.67 ± 5.56		.11 (− .33, .55)	.92 (.88, .95)
Tobacco					
Used *n* (%)	98 (96.1%)	96 (95%)	.88		
Units used (cig) ± SD	20.07 ± 12.51	19.79 ± 13.13		.28 (− 1.26, 1.81)	.81 (.74, .87)
Days used ± SD	26.9 ± 5.46	26.61 ± 6.10		.28 (− .27, .83)	.89 (.83, .92)
Injecting drug use					
Days inject. use ± SD	1.42 ± 5.57	1.24 ± 5.21		.19 (− .00, .37)	.98 (.98, .99)
Shared inj. use n (%)	1(1%)	2(2%)	NC[e]		
Days paid work ± SD	7.40 ± 10.55	7.00 ± 10.03		.40 (− .54, 1.34)	.89 (.84, .93)
Days in school ± SD	.82 ± 3.78	1.01 ± 3.82		− .19 (− 1.14, .77)	.18 (− .02, .36)
Days sports ± SD	2.75 ± 6.91	3.83 ± 8.03		− 1.09 (− 2.88, .71)	.26 (.07, .43)
Homeless *n* (%)	16 (15.7%)	16 (15.7%)	.85		
Eviction *n* (%)	9 (8.9%)	13 (12.7%)	.39		
Care for child > 6 *n* (%)	18 (18.2%)	16 (15.7%)	.82		
Care for child < 5 *n* (%)	16 (15.7%)	17 (17%)	.82		
Arrest *n* (%)	13 (12.7%)	14 (13.7%)	.79		
Violence perp. *n* (%)	28 (27.5%)	27 (26.5%)	.83		

Table 2 (continued)

HTOP item	Test	Retest	K[a]	Mean diff. (95% CI)	ICC[b] (95% CI[c])
Violence victim *n* (%)	32 (31.4%)	33 (32.4%)	.57		
Psychol. health ± SD	4.82 ± 2.46	4.73 ± 2.63		.98 (− .33, .53)	.63 (.50, .74)
Physical health ± SD	5.69 ± 2.67	5.43 ± 2.90		.36 (− .07, .79)	.69 (.58, .78)
Quality of life ± SD	5.58 ± 2.50	5.29 ± 2.42		.28 (− .10, .67)	.68 (.56,.77)

[a]*k* Cohen's kappa,[b]*ICC* intraclass correlation coefficient, [c]*CI* confidence intervals, [d]*SD* standard deviation, [e]*NC* due to the low number of cases, it was not possible to calculate statistics

Table 3 Concurrent validity of HTOP with WHOQOL-BREF/SF-36/AUDIT/EuropASI ($N = 147$)

HTOP item	Validation item	Spearman's *rho*[a]
Psychological health	WHOQOL Psychological Health	0.56
Psychological health	SF36 Mental Health Scale	0.62
Physical health	WHOQOL Physical Health	0.48
Physical health	SF36 Physical Health Scale	0.52
Quality of life	WHOQOL Global quality of life Item	0.66
Quality of life	WHOQOL Physical Health	0.48
Quality of life	WHOQOL Psychological Health	0.61
Quality of life	WHOQOL Social Relationships	0.51
Quality of life	WHOQOL Environment	0.50
Alcohol Days used	AUDIT-c alcohol use days	0.90
Injected use days	EuropASI injected use days	0.93
Shared injected use	EuropASI shared paraphernalia	-

[a] $p < .001$

Table 4 Validity of self-report substance use (SR[a]) with toxicology urine test (UT[b])

HTOP item	Sensitivity	SR + /UT +	Specificity	SR-/UT-	k[c] (95% CI)
Cannabis	0.93	49/45	0.84	39/43	0.77 (0.64, 0.91)
Heroin	0.93	47/40	0.85	61/68	0.75 (0.61, 0.87)
Benzodiazepines	0.96	63/45	0.71	50/68	0.62 (0.48, 0.76)
Cocaine	0.88	24/17	0.88	66/73	0.66 (0.47, 0.84)
Amphetamines	0.40	9/10	0.93	72/71	0.34 (0.02, 0.63)
Other opioids	0.26	16/39	0.91	91/68	0.20 (0.01, 0.36)

[a]*SR* self-report, [b]*UT* urine test, [c]*k* Cohen's kappa

proportion of negative and positive agreement between self report and urine tests for all substances assessed except of amphetamines. A similar finding was reported by Ryan et al. (2014) and may be attributed to the fact that the contrast was underpowered.

However, certain items demonstrated poor inter-rater reliability, namely, opioid and benzodiazepine use, danger of eviction, and days of school/volunteer work. We attribute these findings to patients' unclear definitions for these items, which led to inconsistencies

in reporting relevant behaviors. For example, abuse of illegally obtained extra doses of buprenorphine, methadone, or opioid painkillers is not perceived as "other opioid use"; partner or parents threatening to change the locks is often conceptualized as "danger of eviction." To address this, we further clarified these definitions in the quick reference guide, so as to incorporate them in the typical administration instructions.

Moving forward, testing the instrument with the revised guidelines for the poorer performing items would be worthwhile. Also, it would be useful to examine the HTOP's psychometric properties in different samples. For example, its psychometric properties could be tested in more patients that receive treatment with methadone, in those that attend non-pharmacological therapeutic programs for substance dependence, in adolescents, at programs in other regions of Greece, and also, in psychiatric populations, since it could be very helpful as a screening measure for substance abuse.

In any case, the HTOP is an extremely useful instrument that assesses the patients' therapeutic course briefly and comprehensively, and it may contribute to a more targeted and individualized patient care and a system that refers patients to services that correspond more efficiently to their changing therapeutic needs.

The cultural adaptation of the HTOP and its employment in the Greek context is expected to provide better feedback to both patients and therapists concerning the therapeutic progress in the course of time, to inform the therapeutic planning through patient assessment on substance use and other aspects of their health and social life, to advance the communication between therapists and services during data collection and service documentation, to evaluate therapeutic services through clinical outcome assessment, and also to improve the quality of existing services and design new ones.

Limitations of the Study

The current study does not come without limitations. As mentioned above, all the participants are adult patients attending opioid substitution therapy programs in OKANA in Athens; therefore, the HTOP needs to be further examined in broader treatment populations such as persons who are using different treatment modalities (i.e., withdrawal and detox services), persons residing in rural areas or prisoners attending addiction therapy programs.

Moreover, the present study did not assess change sensitivity, i.e., the ability of the questionnaire to capture changes that occur over time, regardless of whether they are due to treatment or not, by re-administering the tool after a sufficient period of time (Guyatt et al., 1987). Furthermore, questionnaire's construct validity was not assessed. Future research is needed to fill this gap.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11469-021-00630-y.

Acknowledgements The authors would like to thank Paraskevi Psichogiou for her contribution with the data entry.

Author Contribution Conceptualization: Paraskevi Karakoula, Konstantinos Kokkolis, Charikleia Tsatsaroni, Aggeliki Kitsiou, Nikolaos Lintzeris. Data curation: Paraskevi Karakoula, Charikleia Tsatsaroni, Aggeliki Kitsiou. Formal analysis: Paraskevi Karakoula, Sofia Nikolaidi. Investigation: Paraskevi Karakoula, Konstantinos Kokkolis, Charikleia Tsatsaroni, Aggeliki Kitsiou, Emmanouil Koukidis, Paraskevi Danousi, Georgios Kolios, Georgia Vlachou, Apostolia Fragkouli. Methodology: Paraskevi Karakoula, Charikleia Tsatsaroni, Aggeliki Kitsiou, Apostolia Fragkouli, Nikolaos Lintzeris. Project administration: Paraskevi Karakoula, Charikleia Tsatsaroni, Aggeliki Kitsiou. Resources: Konstantinos Kokkolis. Supervision: Nikolaos Lintzeris. Validation: Sofia Nikolaidi. Visualization: Paraskevi Karakoula, Konstantinos

Kokkolis. Writing-original draft: Paraskevi Karakoula, Konstantinos Kokkolis. Writing — review and editing: Konstantinos Kokkolis, Charikleia Tsatsaroni, Nikolaos Lintzeris. All authors approve the final version of the manuscript and agree to be accountable for all aspects of the work.

Data Availability Data are available from the corresponding author upon request**.**

Declarations

Ethics Approval Research was approved by the review board of OKANA. All procedures followed were in accordance with the ethical standards of the responsible committee on human experimentation (institutional and national) and with the Helsinki Declaration of 1975, as revised in 2000 (5).

Consent to Participate Informed consent was obtained from all participants.

Conflict of Interest The authors declare no competing interests.

References

Affholter, D. P. (1994). Outcome monitoring. In: Wholey J. S., Hatry H. P., Newcomer K. E., editors. *Handbook of practical program evaluation 3rd edition*. San Francisco: Jossey-Bass.

Anagnostopoulos, F., Niakas, D. A., & Pappa, E. (2005). Construct validation of the Greek SF-36 health survey. *Quality of Life Research, 14*(8), 1959–1965. https://doi.org/10.1007/s11136-005-3866-8

Bush, K., Kivlahan, D.R., McDonell, M.B., Fihn, S.D. & Bradley, K. A. (1998). The AUDIT alcohol consumption questions (AUDIT-C): An effective brief screening test for problem drinking. Ambulatory Care Quality Improvement Project (ACQUIP). Alcohol Use Disorders Identification Test. *Archives of Internal Medicine 158*(16), 1789–1795.

Cacciola, J. S., Alterman, A. I., DePhilippis, D., Drapkin, M. L., Valadez, C., Fala, N. C., Oslin, D., & McKay, J. R. (2013). Development and initial evaluation of the brief addiction monitor (BAM). *Journal of Substance Abuse Treatment, 44*(3), 256–263.

Castillo-Carniglia, Á., Marín, J. D., Soto-Brandt, G., Donoso, M. P., Piñol, D., San Martín, J., Huepe, D., Alvarado, R., Eastwood, B., & Portilla Huidobro, R. (2015). Adaptation and validation of the instrument treatment outcomes profile to the Chilean population. *Journal of Substance Abuse Treatment, 56*, 39–47.

Cicchetti, D. V., & Sparrow, S. S. (1981). Developing criteria for establishing interrater reliability of specific items: Applications to assessment of adaptive behavior. *American Journal of Mental Deficiency, 86*, 127–137.Dancey, C., & Reidy, J. (2007). *Statistics without Maths for Psychology*. Pearson Education UK.

Cicchetti, D. V., & Feinstein, A. R. (1990). High agreement but low kappa: II. Resolving the paradoxes. *Journal of Clinical Epidemiology, 43*, 551–558.

Deacon, R. M., Mammen, K., Holmes, J., Dunlop, A., & Bruno, r., Mills, L., Graham, R., & Lintzeris, N. . (2020). Assessing the validity of the Australian treatment outcomes profile for telephone administration in drug health treatment populations. *Drug and Alcohol Review, 39*(5), 441–446.

EuropASI Working Group. (1994). *European addiction severity index EuropASI, a guide to training and administering EuropASI interviews, COST A6.* Retrieved August 9, 2019, from: http://www.emcdda.europa.eu/html.cfm/index3647EN.html

Fischer, B., Powis, J., Firestone, C. M., Rudzinski, K., & Rehm, J. (2008). Hepatitis C virus transmission among oral crack users: Viral detection on crack paraphernalia. *European Journal of Gastroenterology and HepatolOgy, 20*(1), 29–32. https://doi.org/10.1097/MEG.0b013e3282f16a8c

Ginieri-Coccossis, M., Triantafillou, E., Tomaras, V., Soldatos, C., Mavreas, V., & Christodoulou, G. (2012). Psychometric properties of WHOQOL-BREF in clinical and health Greek populations: Incorporating new culture-relevant items. *Psychiatriki, 23*(2), 130–142.

Government Gazette. *Law 4139/2013. On substance abuse and other provisions*. (In Greek) Retrieved from: http://www.et.gr/idocs-nph/search/pdfViewerForm.html?args=5C7QrtC22wEaosRGzKxO6XdtvSoClrL8R5ab1UmX_HjtIl9LGdkF53UIxsx942CdyqxSQYNuqAGCF0IfB9HI6qSYtMQEkEHLwnFqmgJSA5WIsluV-nRwO1oKqSe4BlOTSpEWYhszF8P8UqWb_zFijC919zh1PXYUbKSujmUDL7aGixWmIwC5RhUCjeETtoAi

Guyatt, G., Walter, S., & Norman, G. (1987). Measuring change over time: Assessing the usefulness of evaluative instruments. *Journal of Chronic Diseases, 40*(2), 171–178. https://doi.org/10.1016/0021-9681(87)90069-5

Hser, Y. I., Evans, E., Huang, D., & Anglin, D. M. (2004). Relationship between drug treatment services, retention, and outcomes. *Psychiatric Services, 55*, 767–774.

Hunt, P., Hunter, S. B., & Levan, D. (2017). Continuous quality improvement in substance abuse treatment facilities: How much does it cost? *Journal of Substance Abuse Treatment, 77*, 133–140.

Kokkevi, A., & Hartgers, C. (1995). EuropASI: European adaptation of a multidimensional assessment instrument for drug and alcohol dependence. *European Addiction Research, 1*(4), 208–210.

Kraemer, H. C., & Thiemann, S. (1987). *How many subjects? Statistical power analysis in research*. Sage Publications.

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

Lawrinson, P., Gerber, S., Copeland, J., & Indig, D. (2003). *The brief treatment outcome measure: Detoxification, rehabilitation and counselling manual*. National Drug and Alcohol Research Centre.

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.

Leonhard, C., Mulvey, K., Gastfriend, D. R., & Shwartz, M. (2000). Addiction Severity Index: A field study of internal consistency and validity. *Journal of Substance Abuse Treatment, 18*(2), 129–135.

Lintzeris, N., Monds, L., Rivas, G., Leung, S., Withall, A., & Draper, B. (2016). The australian treatment outcomes profile instrument as a clinical tool for older alcohol and other drug clients: A validation study. *Drug and Alcohol Review, 35*(6), 673–677.

Luty, J., Varughese, S., & Easow, J. (2009). Criminally invalid: The treatment outcome profile form for substance misuse. *Psychiatric Bulletin, 33*(11), 404–406.

Malliori, M., Golna, C., Souliotis, K., & Hatzakis, A. (2013). Financial crisis, austerity, and health in Europe. *Lancet, 382*(9890), 392. https://doi.org/10.1016/S0140-6736

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

McLellan, A. T., Kushner, H., Metzger, D., Peters, R., Smith, I., Grissom, G., Pettinati, H., & Argeriou, M. (1992). *Addiction Severity Index-Fifth Edition, 9*(3), 199–213.

Mills, L., Lintzeris, N., Bruno, R., Montebello, M., Dunlop, A., Deacon, R. M., Copeland, J., Jefferies, M., Rivasi, C., & Mammen, K. (2020). Validation of the Australian treatment outcomes profile for use in clients with cannabis dependence. *Drug and Alcohol Review, 39*(4), 356–364. https://doi.org/10.1111/dar.13050

Pappa, E., Kontodimopoulos, N., & Niakas, D. (2005). Validating and norming of the Greek SF-36 Health Survey. *Quality of Life Research, 14*(5), 1433–1438.

Paraskevis, D., Nikolopoulos, G., Fotiou, A., Tsiara, C., Paraskeva, D., Sypsa, V., & Hatzakis, A. (2013). Economic recession and emergence of an HIV-1 outbreak among drug injectors in Athens metropolitan area: A longitudinal study. *PLoS ONE, 8*(11), e78941. https://doi.org/10.1371/journal.pone.0078941

Rosen, C. S., Henson, B. R., Finney, J. W., & Moos, R. H. (2000). Consistency of self-administered and interview-based addiction severity index composite scores. *Addiction, 95*(3), 419–425.

Ryan, A., Holmes, J., Hunt, V., Dunlop, A., Mammen, K., Holland, R., Sutton, Y., Sindhusake, D., Rivas, G., & Lintzeris, N. (2014). Validation and implementation of the Australian treatment outcomes profile in specialist drug and alcohol settings. *Drug and Alcohol Review, 33*(1), 33–42.

Saunders, J. B., & Aasland, O. G. (1987). WHO collaborative project on the identification and treatment of persons with harmful alcohol consumption. *Report on phase I: development of a screening instrument*. World Health Organization, Geneva.

Simpson, D. D. (1979). The relation of time spent in drug abuse treatment to posttreatment outcome. *The American Journal of Psychiatry, 136*(11), 1449–1453. https://doi.org/10.1176/ajp.136.11.1449

Simpson, D. D. (1981). Treatment for drug abuse: Follow-up outcomes and length of time spent. *Archives of General Psychiatry, 38*, 875–880. https://doi.org/10.1001/archpsyc.1981.01780330033003

Teruya, C., Hardy, M., Hser, Y., & Evans, E. (2006). Implementation of a statewide outcome monitoring system: Lessons learned from substance abuse treatment provider staff. *Qualitative Health Research, 16*(3), 337–352.

Thomaidis, N. S., Gago-Ferrero, P., Ort, C., Maragou, N. C., Alygizakis, N. A., Borova, V. L., & Dasenaki, M. (2016). Reflection of socio-economic changes in wastewater: Licit and illicit drug use patterns.

Environmental Science and Technology, 50(18), 10065–10072. https://doi.org/10.1021/acs.est.6b02417

University Mental Health Research Institute and National Documentation and Information Center for Drugs and Drug Addiction (EPIPSI & EKTEPN). (2018). *Problematic drug and alcohol consumption in Greece, Annual Report 2017*. Author.

University Mental Health Research Institute and National Documentation and Information Center for Drugs and Drug Addiction (EPIPSI & EKTEPN) (1996). *European index for the severity of addiction (EuropASI): EuropASI training interview guide* (Ed., Trans., EPIPSI). Athens: Author.

Viera, A. J., & Garrett, J. M. (2005). Understanding interobserver agreement: The kappa statistic. *Family Medicine, 37*(5), 360–363.

Wang, M., Shen, J., Liu, X., Deng, Y., Li, J., Finch, E., & Wolff, K. (2017). Reliability and validity of the treatment outcome profile among patients attending methadone maintenance treatment programs in Kunming, China. *Journal of Substance Abuse Treatment, 77*, 89–94.

Ware, J.E. Jr., & Sherbourne, C.D. (1992). The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Medical Care, 30*(6), 473–483.

WHOQOL Group. (1994). Development of the WHOQOL: Rationale and current status. *International Journal of Mental Health, 23*, 24–56.

WHOQOL Group. (1995). The World Health Organization quality of life assessment (WHOQOL): Position paper from the World Health Organization. *Social Science and Medicine, 41*, 1403–1409.

WHOQOL Group. (1998). Development of the World Health Organization WHOQOL-BREF Quality of life assessment. *Psychological Medicine, 28*, 551–558.

WHOQOL Group (1994b). The development of the World Health Organization quality of life assessment instrument (the WHOQOL). In: Orley J, Kuyken W (eds), *Quality of life assessment: International perspectives*, Heidelberg: Springer-Verlag, 41-57