

PLEASE NOTE:

This is the author's version of the manuscript accepted for publication in *Drug and Alcohol Dependence*. Changes resulting from the publishing process, namely editing, corrections, final formatting for printed or online publication, and other modifications resulting from quality control procedures, may have been subsequently added.

The published version can be found in: Alves, P., Sales, C. M. D., & Ashworth, M. (2017). Does outcome measurement of treatment for substance use disorder reflect the personal concerns of patients?: A scoping review of measures recommended in Europe. *Drug and Alcohol Dependence*, 179, 299-308. doi: 10.1016/j.drugalcdep.2017.05.049

Does outcome measurement of treatment for substance use disorder reflect the personal concerns of patients? A scoping review of measures recommended in Europe

Paula Alves¹, Célia Sales², Mark Ashworth³

¹. University College London; London, United Kingdom

². Universidade do Porto; Porto, Portugal

³. King's College London; London, United Kingdom

Correspondence:

Paula Alves

paulagomesalves@hotmail.com

Abstract

There are a growing number of authors stating that outcome measurement in treatment for substance use disorders should go beyond substance use and include other bio-psycho-social variables of interest. However, little is known about which topics tend to be covered by outcome measures and whether they reflect the typical concerns of this patient group. This study followed a scoping review methodology in which 42 outcome measures recommended by an EU agency for substance use disorders were reviewed. We identified the domains of problems covered by these 42 measures and then compared them with 54 domains derived from patients, in a previous study. We also explored how similar the existing measures were in terms of domains covered, and which patient derived domains tended to be represented in those measures. We identified 31 domains of problems across the 42 measures, with ‘substance use’ and ‘psychological health’ among the commonest. Most measures were similar in content to each other and multidimensional. Almost all domains of problems identified in the outcome measures corresponded to concerns reported by patients. On the other hand, we found that several topics of relevance for patients were not covered by any of the measure included in our study. This suggests that existing outcome measurement does not always target aspects that affect patients’ lives, as reported directly by patients. Our study shows that outcome measurement needs to adopt a more flexible and comprehensive approach, by taking on board the problems experienced by patients in this population.

Key words: Outcome measurement; patient perspectives; thematic analysis.

1.0 Introduction

In substance use disorders, as in any other mental health intervention, outcome measurement plays a crucial role in understanding whether interventions are effective. Little is known about whether the traditional outcome measures used in this population cover the problems experienced by patients. This study aims to take a step in this direction, by reviewing outcome measures that are used to evaluate treatment for substance use disorders; as well as comparing the types of problems that such measures cover with the problems that patients in this population report having, when asked to talk freely about their clinical situation.

1.1 Outcome Measurement In Large-Scale Studies

To date, one of the largest outcome studies in this field was the National Treatment Outcome Research Study (NTORS; Gossop et al., 2003), conducted in the United Kingdom in the early 2000s. This study investigated the outcomes of community and residential treatment in a national sample of 418 individuals using illicit drugs and alcohol. It revealed reductions in drug use and risk-taking behaviours, as well as improvements in psychological and physical health. Other large-scale outcome studies have reported similar findings (e.g., ATOS; Darke et al., 2007; DTORS; Jones et al., 2009). Gossop (2008) also referred to a “clinical fallacy”, which considers that numerous successful cases are overlooked in reports due to patients leaving the services voluntarily after self-perceived improvements. Although treatment drop-out can also represent treatment failure, these findings indicate that treatment for substance use disorders works, at least to a certain extent, and for a certain group of patients.

However, for a relatively large proportion of people, treatment for substance use disorders is not as successful as desired. For instance, in NTORS, after a 5-year follow-up, there was still a 1% mortality rate and the levels of alcohol consumption remained unchanged

among those that took part in the study. Moreover, an increase in frequency of alcohol use was observed in patients receiving residential treatment (Gossop, 2003). The latest report published by the National Drug Treatment Monitoring System of Public Health England (PHE, 2015) showed that approximately one third of patients (with opiate, non-opiate, non-opiate/alcohol or alcohol problems only) left treatment prior to its completion or against the clinical judgement of practitioners (i.e., drop-out). It also revealed that one third of patients with opiate and/or non-opiate dependence problems remain unchanged after six months in treatment; and considering solely the alcohol problems, nearly half of all patients failed to improve.

The reasons for these discouraging results are unclear, but unsuccessful treatment may negatively affect patients' mental, physical and social status, and contribute to mortality excess in this group. On the other hand, individuals may have a variety of concerns or health problems that is not captured by traditional methods of outcome measurement. Hence, there is a need to gain a better sense of which outcomes are relevant for this patient group and to reflect upon the value and relevance of the information which is being used for outcome measurement purposes.

1.2 How Should We Measure The Outcomes Of Treatment?

In 2011, Tiffany and colleagues stated that, to be effective, the measurement of treatment outcomes for substance use disorders should 1) focus on the consequences or strong, concurrent correlates of excessive drug use; 2) cover aspects that are common across abused substances and widespread among people dependent on those substances; and 3) have documented and strong psychometric properties that produce replicable evidence that the outcome can be altered following treatment. Despite providing a good framework, or strategy, for the evaluation of treatment for substance use disorders, Tiffany and colleagues

(2011) proposal does not suggest which aspects should be covered by outcome measures in this field.

There is a great diversity of outcome criteria proposed for substance use disorders treatment. However, the literature lacks consensus, with many international organisations (e.g., European Monitoring Centre for Drugs and Drug Addiction, EMCDDA) and authors suggesting different domains, or outcome measurement criteria (Table 1).

Moreover, there is a gap between what is proposed in the literature and what is actually evaluated in research studies and/or practice. Most evaluation protocols focus on drug and/or alcohol use and related behaviours, e.g., injecting, criminal activities (Donavan et al., 2011; Tiffany et al., 2011), overlooking psychosocial variables that many authors believe to be highly relevant for patients' recovery (Table 1). These data could be used as markers to adjust the intervention according to treatment response (Tiffany et al., 2011), as well as allowing for a better understanding of the mechanisms underlying recovery.

1.3 Does Outcome Measurement Reflect The Concerns Of Patients?

The outcome measurement process in this patient group faces another major challenge: most measures are expert-driven, and not primarily sourced from patient perspectives. Failure to involve patients in the process of outcome measurement raises the possibility that measurement overlooks aspects of relevance for patients (Alves et al., 2016) and over-optimistic reporting of outcomes (Thurgood et al., 2014). Increasingly, studies are seeking the views of patients about outcome measurement criteria (e.g., Ruefli and Rogers, 2004; Neale et al., 2016; see Table 1). For instance, Ruefli and Rogers (2004) revealed that patients in treatment stated the importance of domains covering: 'making money', 'getting something good to eat', 'being housed', 'relating to family', 'getting needed programs/benefits/services', 'handling health problems', 'handling negative emotions', 'handling legal problems', 'improving oneself' and 'handling drug-use problems'. In a

similar study conducted in 2015, Neale and colleagues asked patients to define what “being recovered” meant for them. Patients reported that treatment recovery should include improvements in ‘substance use’, ‘material resources’, ‘outlook on life’, ‘self-care’ and ‘relationships’. Another example is the recently developed ‘SURE’, a standardised outcome measure for treatment of substance use disorders (Neale et al., 2016). In this measure, items were generated in collaboration with former and current drug and alcohol service users (Neale et al., 2016). Similarly, our own research group has sought the views of patients with substance use disorders using individualised outcome measures, which enable patients to report their personal concerns at treatment entry (Alves et al., submitted). We found that patient priorities were ‘addictive behaviour’, ‘work-related problems’, ‘general relationships with family’, ‘money’ and ‘worries about another person’. Additionally, we found that most of the patient-reported topics were not captured by widely used standardised outcome measures of psychological well-being, such as CORE-OM or PHQ-9, or drug-related outcome measures, as TOP (Alves et al., submitted). These findings suggest that patients may contribute new evidence to outcome measurement. Such evidence may provide insight into contradictory findings reported in the literature about the outcomes of treatment for substance use disorders.

1.4 Study Rationale

The principal aim of this study was to explore the thematic content of items used by outcome measurement tools in treatment for substance use disorders and to understand the extent to which these measures are perceived as relevant to this population. We hope that our findings will contribute to a broader understanding of the attributes of outcome measures and their ability to capture the personalised problems reported by patients.

2.0 Methods

This study was comprised of three analytical steps: to search for outcome measures and to identify the domains covered by those measures; to explore the extent to which the domains were duplicated in different measures, or were unique; and finally, to conduct a thematic comparison analysis comparing the ‘problem’ domains in standardised instruments with domains generated by patients, which had previously been derived from individualised measures in a previous study (Alves et al., submitted).

2.1 Search Strategy, Selection Of Measures And Data Extraction

A scoping review approach (Arksey and O’Malley, 2002) was used to extract information from the outcome measures. This approach was considered as the most appropriate because it allows for a rapid mapping of key concepts in a certain field, from the most relevant/main sources. With this methodology in mind, the key concepts used to chart our data were the general characteristics and domains covered by the outcome measures.

For this study, we restricted our search to outcome measures used in Europe to match the setting in which the patient-driven domains used for comparison (see Data Analysis section below) were developed. We believed this matching would decrease the existence of confounding factors (e.g., cultural differences) potentially affecting this comparison. After selecting the European setting, we started by hand searching for outcome measures at the website of EMCDDA, an EU Agency that aims to provide “factual, objective, reliable and comparable information concerning (...) drug addiction and their consequences” (see <http://www.emcdda.europa.eu/about>), including public policies, health responses and treatment. This Agency collects data about substance use disorder across 30 European countries. Stored data can be retrieved from the ‘Evaluation Instrument Bank’ (<http://www.emcdda.europa.eu/eib>). Twenty-six national agencies for drug and alcohol monitoring were also contacted by e-mail to ensure that all measures recommended for

outcome measurement had been identified. No additional tools were identified by the six national agencies that responded to our request.

Outcome measures were selected for review according to two inclusion criteria: 1) used to evaluate the outcome of treatment; and 2) used in adults as the target population. The measures were excluded if: 1) they did not target the patient (e.g., measures focusing on significant others); and 2) were not available in English. After selecting the outcome measures, a charting form was created for data recording. We extracted data describing general characteristics of the measures, by identifying their authors, year of publication, focus (drugs, alcohol, drug/alcohol, health, other), type of measure (standardised, individualised, hybrid), delivery format (self-report, interview, mixed) and number/type of items (Likert scales, nominal scales, number of days/times, other).

2.2 Data Analysis

2.2.1 Part 1: Identification of domains covered by outcome measures. We started by reading the Evaluation Instruments Bank's "User information" sheet of each measure, which included a section about "Domains measure/life areas/problems assessed". If unavailable, the accompanying instructions or main references were consulted. A preliminary list of domains was then created, containing all domains included in each selected measure. In this list, domains were recorded *verbatim* as provided in the description of the measure. We removed duplicate domains and grouped those that referred to similar/equivalent topics, based on linguistic proximity (i.e., synonyms) until a complete list of domains was obtained. For instance, 'illegal activities' and 'criminal involvement' were both integrated in the domain 'crime'. In the case of 'psychological health', we opted for creating several sub-domains, due to the existence of instruments that either focussed on general or single aspects of psychological health (e.g., 'self-esteem'). This procedure was performed by the first author

and the results were discussed with a researcher, independent to the study, until consensus was reached.

2.2.2 Part 2: Similarity between outcome measures in terms of domains covered.

Next, we grouped the measures according to the similarity of the domains which they contained, to understand how convergent the tools were. To achieve this goal, we used the Metric-Frequency similarity index (Sales and Wakker, 2009; Sales et al., 2015) to obtain a similarity matrix, based on the ‘absence’ or ‘presence’ of each domain, in each tool. The similarity matrix was analysed with the Hierarchical Cluster Technique (centroid method) to obtain groups of measures that shared common features. We used IBM SPSS Statistics 22 and the freeware online MF calculator (<http://mfcalculator.celiasales.org/>) to conduct this analysis.

2.2.3 Part 3: Matching between outcome measure domains and the problems relevant for patients. The third part of our analysis comprised a thematic comparison between the domains extracted from the outcome measures and 54 previously defined domains of patient-generated problems. This list of 54 patient-generated domains was identified in a study with 93 patients entering treatment four clinical services for substance use disorders in a previous study (Alves et al., submitted). This study was carried out in three outpatient and one inpatient service for treatment of substance use disorders in Portugal. It included 55 males and 38 females, with a mean age of 42.7 years old ($SD = 11.3$). Nearly half of the sample was married (48%) and the majority received education until years 5-9. More than half of the sample was unemployed (58%). In this study, patients were asked to identify their problems in two individualised outcome measures (one orally administered and one in a pen-and-paper format), as they were entering their first treatment session. Individualised measures are tools where individuals can report the concerns that led them to treatment, in an open-ended

fashion and in their own words. The domains of patient-generated problems were derived from the categorisation of these problems, according to their underlying sub-theme.

The thematic comparison was made independently by two researchers with post-graduate education in psychology and trained with the individualised outcome measures and the coding system used in this study. The researchers used a binary coding system to rate each outcome measure domain as “yes” (i.e., topics connected, clearly related or completely overlapped with patient-generated domains) and “no” (i.e., topics completely different from patient-generated domains). This procedure was followed by calculations of inter-rater reliability, based on Cohen’s kappa values. Certain outcome measures selected for review contained sections about socio-demographic/treatment history and treatment process (e.g., readiness for treatment). These sections were excluded from our thematic comparison, as we were interested in outcome criteria only. After completing the content matching, we calculated the number and percentage of patient-generated domains included in each of the outcome measures.

Finally, for data reduction purposes, we re-analysed the data obtained from the thematic comparison of patient concerns to identify the presence of underlying constructs featured in the selected outcome tools. We only included those patient-generated domains which were identified in at least 10% of the outcome measures. This cut-off was selected because less frequently featured domains were not present in a wide range of measures. This produced a cohort of 20 domains; the remaining 34 domains were discarded from this part of the analysis. To explore the structure of underlying constructs in selected domains, we used multiple correspondence analysis (MCA; Carvalho, 2008). MCA is a multivariate technique used to detect underlying dimensions among a group of nominal/categorical data. Factor analysis can also be used for a similar purpose; however, unlike factor analysis, MCA makes no assumptions of normality and can be used with categorical data (Carvalho, 2008; Philips,

2009). Using MCA, dimensions are identified on the basis of their discriminatory and contribution values, in a range varying from zero to one.

3.0 Results

After removing the duplicates, 74 outcome measures were screened in this study, of which 42 fulfilled the selection criteria (see selection flowchart in Figure 1).

3.1 Main Characteristics Of The Outcome Measures

Among the 42 outcome measures, 25 (60%) focussed on drugs/alcohol, seven (17%) focused on drugs, six (14%) on alcohol, one (2%) on general health and three (7%) on other aspects (i.e., depression, self-esteem and quality of life). All but one measure were standardised, and only one differed by having an individualised scope, i.e., non-pre-set defined items. On delivery format, 21 (50%) were interview-based protocols, 19 (46%) were self-report measures, one (2%) was an observational scale and one (2%) followed a mixed-methods approach. The mean number of items was 54 (SD = 57, range 1-223). Among these measures, 23 (55%) contained nominal items and 34 (81%) had scale-type items (see Table 2).

3.2 Domains Identified And Similarity Between Outcome Measures According To Domains Covered

From a preliminary group of 196 domains of problems identified across the measures, we obtained a final list of 31 domains by excluding all domains which had overlapping content. The full list of these domains is presented in Figure 2.

The mean number of domains per measure was 4 (SD = 4, range 1-18); and 43% of measures (n = 18) covered only one domain. The most common domains were: 'substance use' (67%), 'psychological health general' (40%), 'family and social relationships' (38%), 'employment' (38%), 'crime' (36%) and 'physical health' (26%).

When exploring the similarity between the instruments in terms of domains covered, we found two main groups of measures: those that focussed only on substance use; and multidimensional measures focusing on substance use and other topics. A similarity tree representing these groups is shown in Figure 3.

3.3 Comparison between outcome measure domains and patient-relevant problems

The content matching between the 31 domains and the 54 domains of patient-generated problems was performed with satisfactory inter-rater agreement values (Cohen's *kappa* between raters ranged from 0.65 to 1.0). Overall, most patient-generated domains ($n = 34$, 63%) were represented by a small proportion (10%) of outcome measures (see Table 3). Moreover, 26% of patient-generated domains ($n = 14$) did not feature in any of the outcome measures. Among the unmatched patient-generated domains were topics such as 'personal development', 'understanding self', 'existence', 'outlook on life', 'moving on', 'guilt', 'bereavement' and 'dependence on other people'. On the other hand, nearly all 31 domains found in the outcome measures were matched, in terms of topics, to at least one patient-generated sub-theme. The exceptions were the domains: 'domestic life', 'self-care', 'daytime activities', 'technology and information' and 'acculturation'. We also found that the patient-generated domains of problems frequently represented in outcome measures tend to focus on four main areas: interpersonal relationships (dimension 1); communication problems (dimension 2); addiction (dimension 3); and social problems (dimension 4) (see Table 4).

4.0 Discussion

This study provides an overview of the measures recommended for outcome measurement in treatment for substance use disorders provided in Europe. Our goal was to identify the main characteristics of these measures; and, to explore the extent to which they covered individualised problems that bring patients to substance use disorders treatment. In the next paragraphs, we discuss the lessons derived from our findings.

4.1 Lesson 1: Most outcome measures evaluate similar domains and are predominantly multidimensional

In the 1960s professionals felt the need to employ methods to quantify substance use disorders that were being overlooked, which resulted in the proliferation of outcome tools (Dwyer and Fraser, 2015). Our study reveals that there is considerable duplication of the contents of outcome measures with consequent redundancy of measures in terms of topics covered. Hence, careful consideration of the domains worth measuring is required before selecting an outcome measure (Slade and Thornicroft, 2014).

Our study revealed that outcome measures used in substance use disorders treatment can be clustered into two main groups. One group included unidimensional measures focussing on substance use. This was consistent with the study population and with the fact that most outcome studies focus solely on this topic. A second and larger group contained three types of measures covering: substance use and behaviours associated with this disorder (e.g., injection of drugs); substance use and psychological health; and domains not directly related to substance use (e.g., self-esteem). Psychological health was the second most commonly identified domain, featuring in nearly half the measures. Given the importance of psychological health, these findings suggest that outcome measurement currently adopts a strategy to measure patients' changes from a psychological point of view.

These findings show that the “narrow” approach (Bühringer, 2012) that most outcome studies adopt is not due to the lack of multidimensional tools. Hence, it is the paradigm underlying outcome measurement that could adopt a broader approach by using more comprehensive measures. This is consistent with studies where patients in treatment for substance use disorders reported a need to talk about topics that go beyond their drug-related problems (Alves et al., 2016). Ultimately, using multidimensional tools could not only expand the scope of outcome measurement, by covering a greater variety of problems

experienced by patients, but also shed light on situations that may trigger or motivate substance drug use. In other words, the more information one gathers about the individual circumstances that are related to patients' addictive behaviour, the better prepared we will be to reflect on treatment priorities and relapse prevention.

4.2 Lesson 2: The domains covered by outcome measures are relevant for patients

Considering that outcome measures tend to be standardised and expert-driven, we were particularly interested in exploring the extent to which they reflected the personalised problems of patients in treatment for substance use disorders. We found that most domains (e.g., addiction, work-related problems, relationship difficulties, money worries) covered by outcome measures were equivalent to problem domains reported by patients with substance use disorders. Relevance to patients was demonstrated by findings reported in a previous study in which patients had reported that outcome measurement targeting substance use helped them to think about the negative consequences of their addictive behaviour (Alves et al., 2016). Moreover, the patient-generated domains which also featured in the outcome tools tended to focus on interpersonal and/or social relationships, concerns which are reported by 20% of patients in treatment for substance use disorders, reinforcing the evidence that outcome measures was addressing concepts of interest to patients (Alves et al., submitted).

4.3 Lesson 3: The outcome measures available are overlooking areas of concern to this population

To this point, we have showed that the existing measures tend to cover topics of relevance for patients, which was a positive finding. However, we also found that despite including relevant topics, they tend to miss areas of problems that patients deem as important and that should have been measured. More specifically, our study revealed that only 10% of outcome measures contained half or more of the patient-generated domains identified by our study. The remaining 90% of measures contained under half these domains. Some patient-

generated domains overlooked by outcome measures reflected specific concerns likely to be experienced by a small minority (e.g., ‘bereavement’). However, more universally experienced psychological factors or distressing events such as ‘guilt’, ‘dependence of other people’ or ‘housing problems’ were rarely included in outcome measures. None of the measures included worries about the self (e.g., ‘personal development’, ‘existence’, ‘outlook on life’, ‘having time’) even though these were frequently reported in individualised measures (Alves et al., submitted). Some of these patient-generated topics, namely, ‘self-care’ and ‘outlook on life’, have recently been incorporated into a new standardised measure, ‘SURE’, a patient-reported outcome measure for use in substance use (Neale et al., 2016).

Our study suggests that patients seeking treatment for substance use disorders are likely to have a greater diversity of concerns than has previously been acknowledged by experts in the process of designing outcome measures. However, one cannot expect a single outcome measure to be able to address all concerns that patients have about their lives. For this reason, we suggest the use of a more open-ended approach to outcome measurement which elicits a broader range of information from patients. This open-ended approach would allow patients to talk about whichever topic was troubling them, without limiting the scope of their clinical situation. This could be achieved using the so-called individualised outcome measures (Sales and Alves, 2016), measures that do not contain items created *a priori* by experts, but instead invite patients to write or talk about the main concerns that led them to seek treatment. In addition to informing clinicians about the whole range of patients’ problems, open-ended individualised measures could also identify problems that may be of greater concern for the patient than substance use itself, at least in the short-term, which may result in referral to other specialist services.

4.4 Implications for outcome measurement in treatment for substance use disorders

We believe that our study has provided an insight into the limitations of current outcome measures and how we can improve outcome measurement by producing more informed (and less contradictory) findings about treatment outcomes. Based on our study of measures used in Europe, we recommend that health and policy bodies identify a core-set of outcome criteria for use in treatment evaluation, thus facilitating the selection of outcome tools. However, it is important that both researchers/professionals and patients are engaged with this task. Although standardised outcome measures tend to cover relevant domains, they also overlook relevant information about individual distress. As Slade and Thornicroft (2014) put it, “any attempt to squeeze personal identity into predefined boxes can be justifiably criticised for its loss of meaning” (p. 120). This problem could be overcome using a more patient-centred approach to outcome measurement, by using individualised measures (see Sales and Alves, 2016, for a review about these tools), which allow patients to express their personal problems. We advocate that individualised measures are used in combination with existing standardised measures, which provide population reference data. If the use of individualised measures is not feasible, a standardised patient-reported outcome measure designed with a high level of patient input may be a suitable alternative. Although “imperfect measures are better than no measures at all” (Scheyett et al., 2013), the addition of items deemed important by patients has the potential to improve outcome measurement.

4.5 Limitations

This study is not without limitations. First, we accept that some outcome measures may have been omitted from our study if they were not included in the EMCDDA database. However, our goal was to confine our search to measures recommended by international agencies. Restriction of our search to Europe may have restricted our findings although measures need to be culturally applicable (Mann (2012)). A further limitation is that our review did not differentiate between measures used for subgroups of patients. For instance,

we did not included measures targeting adolescents since the patient-generated domains used for comparison purposes were derived from an adult population. We believe that this present study adds up to an increasing body of literature demonstrating the importance of multidimensional outcome measurement in substance use disorders and the inclusion of patient perspectives. Such an approach would acknowledge the multiplicity of problems associated with substance use disorders, as well as taking into account the concept that recovery is an individual journey.

References

- Alves, P., Sales, C., Ashworth, M., 2016. "It is not just about the alcohol": A focus group to explore service users' views about outcome measurement. *Subst. Abuse Treat. Prev. Policy* 11, 1-7. doi: 10.1186/s13011-016-0070-5.
- Alves, P., Sales, C., Ashworth, M., Faísca, L. "There are things I want to say but you don't ask": A thematic comparison between standardised and individualised outcome measures in substance misuse treatment. Manuscript submitted for publication.
- Annis, H., Martin, G., 1985. *Inventory of Drug-Taking Situations* (4th ed.). Toronto: Addiction Research Foundation.
- Arksey, H., O'Malley, L., 2005. Scoping studies: Towards a methodological framework. *Int. J. Soc. Res. Method.* 8, 19-32. doi: 10.1080/1364557032000119616
- Barber, J.G., Cooper, B.K., Heather, N., 1991. The Situational Confidence Questionnaire (Heroin). *Int. J. Addictions.* 26, 565-575. doi: 10.3109/10826089109058905.
- Bradman, K., Erdmann, A.J., Lorge, I., Wolff, H., 1949. The Cornell Medical Index: An adjunct to medical interview. *JAMA.* 140, 530-534.
- Bühringer, G., 2012. The mountain roared and brought forth a mouse: comments on the results of an expert panel on standardisation of drug dependence treatment trial outcome variables. *Addiction* 107, 722-723. doi: 10.1111/j.1360-0443.2011.03689.x
- Carvalho, H., 2008. *Análise Multivariada de dados qualitativos*. Lisboa: Edições Sílabo.
- Christo, G., Spurrell, S., Alcorn, R., 2000. Validation of the Christo Inventory for Substance-misuse Services (CISS): a simple outcome evaluation tool. *Drug. Alcohol. Depend.* 59, 189-197.
- Darke, S., Ward, J., Hall, W., Heather, N., Wodak, A., 1991. *The Opiate Treatment Index (OTI) manual*. New South Wales: National Drug and Alcohol Research Centre of the University of New South Wales.

- Darke, S., Ross, J., Teesson, M., 2007. The Australian Treatment Outcome Study (ATOS): What have we learnt about treatment for heroin dependence? *Drug Alcohol Rev.* 26, 49-54. doi: 10.1080/09595230601036986
- Donovan, D., Bigelow, G., Brigham, G., Carroll, K., Cohen, A., Gardin, J., Hamilton, J., Huestis, M., Hughes, J., Lindblad, R., Marlatt, A., Preston, K., Selzer, J., Somoza, E., Wakim, P., Wells, E., 2011. Primary outcome indices in illicit drug dependence treatment research: Systematic approach to selection and measurement of drug use end-points in clinical trials. *Addiction* 107, 694-708. doi: 10.1111/j.1360-0443.2011.03473.x
- Dwyer, R., Fraser, S., 2015. Addiction screening and diagnostic tools: ‘Refuting’ and ‘unmasking’ claims to legitimacy. *Int. J. Drug Policy* 26, 1189-1197. doi: 10.1016/j.drugpo.2015.08.016
- Diagnostic Validity of the Drug Abuse Screening Test in the Assessment of DSM-III Drug Disorders.
- Gavin, D., Ross, H., Skinner, H., 1989. Diagnostic Validity of the Drug Abuse Screening Test in the Assessment of DSM-III Drug Disorders. *Br. J. Addict.* 84, 301-307.
- Gossop, M., Darke, S., Griffiths, P., Hando, J., Powis, B., Hall, W., Strang, J. 1995. The Severity of Dependence Scale (SDS): psychometric properties of the SDS in English and Australian samples of heroin, cocaine and amphetamine users. *Addiction*, 90, 607-614. doi: 10.1046/j.1360-0443.1995.9056072.x
- Gossop, M., 2012. The clinical fallacy and treatment outcomes. *Addiction* 103, 89-90.
- Public Health England, 2015. Adult substance misuse statistics from the National Drug Treatment Monitoring System (NDTMS). Public Health England. London, UK.

- Gossop, M., Marsden, J., Stewart, D., Kidd, T., 2003. The National Treatment Outcome Research Study (NTORS): 4-5 year follow-up results. *Addiction*. 98, 291-303. doi: 10.1046/j.1360-0443.2003.00296.
- Handelsman, L., Cochrane, K., Aronson, M., Ness, R., Rubinstein K., Kanof, P., 1987. Two New Rating Scales for Opiate Withdrawal. *Am. J. Drug. Alcohol. Abuse*. 13, 293-308. doi: 10.3109/00952998709001515.
- Jones, A., Donmall, M., Millar, T., Moody, A., Weston, S., Anderson, T., Gittins, M., Abeywardana, V., D'Souza, J., 2009. The Drug Treatment Outcomes Research Study: Final outcomes report. Home Office. London, UK.
- Marsden, J., Gossop, M., Stewart, D., Best, D., Farrell, M., Lehmann, P., Edwards, C., Strang, J., 1998. The Maudsley Addiction Profile (MAP): A brief instrument for assessing treatment outcome. *Addiction*. 93, 1857-1867. doi: 10.1046/j.1360-0443.1998.9312185711.x.
- Martin, G., Wilkinson, D., Poulos, C., 1995. The Drug Avoidance Self-Efficacy Scale. *J. Subst. Abuse*. 7, 151-163. doi: 10.1016/0899-3289(95)90001-2.
- McLellan, A., Kushner, H., Metzger, D., Peters, R., Smith, I., Grissom, G., Pettinati, H., Argeriou, M., 1992. The Fifth Edition of the Addiction Severity Index. *J. Subst. Abuse. Treat.* 9, 199-213.
- Neale, J., Tompkins, C., Wheeler, C., Finch, E., Marsden, J., Mitcheson, L., Rose, D., Wykes, T., Strang, J., 2015. 'You're all going to hate the word 'recovery' by the end of this': Service users' views of measuring addiction recovery'. *Drugs Edu. Prev. Policy* 22, 26-34. doi: 10.3109/09687637.2014.947564
- Neale, J., Vitoratou, S., Finch, E., Lennon, P., Mitcheson, L., Panebianco, D., Rose, D., Strang, J., Wykes, T., Marsden, J., 2016. 'Development and validation of 'SURE': A patient reported outcome measure (PROM) for recovery from drug and alcohol

- dependence'. *Drug Alcohol Depend.* 165, 159-167. doi: 10.1016/j.drugalcdep.2016.06.006
- Radloff, L., 1977. The CES-D scale: A self-report depression scale for research in the general population. *Appl. Psychol. Meas.* 1, 385-401. doi: 10.1177/014662167700100306.
- Raistrick, D., Dunbar, G., Davidson, R., 1983. Development of a questionnaire to measure alcohol dependence. *B. J. Addict.* 78, 89-95.
- Raistrick, D., Bradshaw, J., Tober, G., Weiner, J., Allison, J., Healey, C., 1994. Development of the Leeds Dependence Questionnaire. *Addiction.* 89, 563-572. doi: 10.1111/j.1360-0443.1994.tb03332.x.
- Robinson, S., Sobell, L., Sobell, M., Leo, G., 2012. Reliability of the Timeline Followback for Cocaine, Cannabis, and Cigarette Use. *Psychol. Addict. Behav.* 28, 154-162. doi: 10.1037/a0030992.
- Rosenberg, M. (1965). *Society and The Adolescent Self-Image*. Princeton, N.J.: Princeton University Press.
- Ruefli, T., Rogers, S., 2004. How do drug users define their progress in harm reduction programs? Qualitative research to develop user-generated outcomes. *Harm Reduct. J.* 1, 8-13. doi:10.1186/1477-7517-1-8.
- Sales, C., Wakker, P.P., 2009. The Metric-Frequency Measure of similarity for ill-structured data sets, with an application to family therapy. *Br. J. Math. Stat. Psychol.* 62, 663-682. doi: 10.1348/000711008X376070
- Sales, C., Wakker, P.P., Alves, P., Faísca, L., 2015. MF Calculator: A web-based application for analyzing similarity. *J. Stat. Softw.* 65. doi: 10.18637/jss.v065.c02
- Sales, C., Alves, P., 2016. Patient centred assessment in Psychotherapy: A review of individualised tools. *Clin. Psychol. Sci. Pract.* 23, 265-283. doi:10.1111/cpsp.12162.

- Scheyett, A., DeLuca, J., Morgan, C., 2013. Recovery in severe mental illnesses: A literature review of recovery measures. *Soc. Work Res.* 37, 286-302. doi: 10.1093/swr/svt018
- Schippers, G., Broekman, T., Buchholz, A., Rutten, R., 2009. Introducing a new assessment instrument: The Measurements in the Addictions for Triage and Evaluation. *Sucht.* 55, 209-218. doi: 10.1024/2009.04.03.
- Skevington, S., Lotfy, M., O'Connell, K., WHOQOL Group, 2004. The World Health Organization's WHOQOL-BREF quality of life assessment: psychometric properties and results of the international field trial. *Qual. Life. Res.* 13, 299-310. doi: 10.1023/B:QURE.0000018486.91360.00.
- Skinner, H., Sheu, W., 1982. Reliability of Alcohol Use Indices. The Lifetime Drinking History and the MAST. *J. Stud. Alcohol.* 43, 1157-1170. doi: 10.15288/jsa.1982.43.1157.
- Sobell, L., Sobell, M., 1992. Timeline followback: A technique for assessing self-reported alcohol consumption. In R.Z. Litten & J. Allen (Eds.), *Measuring alcohol consumption: Psychosocial and biological methods* (pp. 41-72). New Jersey: Humana Press.
- Sobell, L., Sobell, M., Buchan, G., Cleland, P. A., Fedoroff, I., Leo, G., 1996. The reliability of the Timeline Followback method applied to drug, cigarette, and cannabis use. Presented at the 30th Annual Meeting of the Association for Advancement of Behavior Therapy. New York, NY.
- Texas Christ University, Institute of Behavioral Research - Forms. (n.d.). Retrieved July 23, 2017, from <https://ibr.tcu.edu/forms/>
- Thurgood, S., Crosby, H., Raistrick, D., Tober, G., 2014. Service user, family and friends' views on the meaning of a 'good outcome' of treatment for an addiction problem. *Drugs Edu. Prev. Policy* 1-9. doi: 10.3109/09687637.2014.899987

- Thornicroft, G., Salde, M., 2014. New trends in assessing the outcomes of mental health interventions. *World Psychiatry* 13, 118-124. doi: 10.1002/wps.20114
- Tiffany, S., Friedman, L., Greenfield, S., Hasin, D., Jackson, R., 2011. Beyond drug use: A systematic consideration of other outcomes in evaluations of treatments for substance use disorders. *Addiction* 107, 709-718. doi: 10.1111/j.1360-0443.2011.03581.x
- Wesson, D., Ling, W., 2003. The Clinical Opiate Withdrawal Scale (COWS). *J. Psychoactive. Drugs.* 35, 253-259. doi: 10.1080/02791072.2003.10400007.
- Winters, K., 1992. Development of an adolescent substance abuse screening questionnaire. *Addict. Behav.* 17, 479-490.

Table 1: *Examples of domains (i.e., outcome criteria) to use during outcome measurement of treatment for substance use disorders, as suggested by experts*

Type of publication	Authors/year	Source	Criteria suggested for outcome measurement
International guidelines	EMCDDA, 2007	EU Agency	Addictive behaviour/consumption of substances, retention/time in treatment, status at discharge (planned / drop-out), risk-taking behaviours for drug-related infection, somatic and psychiatric health, social reintegration (housing, employment, social network, life style, delinquency), and quality of life
Scientific literature	Tiffany et al., 2011	Addiction research experts	Self-efficacy, psychosocial functioning, network/social support, craving and quality of life
Scientific literature	Donavan et al., 2011	Addiction research experts	Behavioural functioning and quality of life
Scientific literature	Neale et al., 2015	Addiction research experts	Substance use, treatment/support, psychological health, physical health, use of time, education/training/employment, income, housing, relationships, social functioning, offending/anti-social behaviour, well-being, identity/self-awareness, goals/aspirations and spirituality

Table 2: *Review of the outcome measures used in treatment for substance use disorders (N = 42), including the identification of domains (N = 31).*

	Author/Year	Focus	Type of measure	Delivery format	No. items	Type of items	Domains
Addiction Severity Index	McLellan et al., 1979	Drugs/alcohol	Standardised	Interview	161	Nominal, scale	Crime, employment, family/relationships, general health, money, psychological health, substance use
Addiction Severity Index Crime	Öberg et al., 1998	Drugs/alcohol	Standardised	Interview	39	Nominal, scale	Crime
Addiction Severity Index Gambling	Öberg et al., 1999	Drugs/alcohol	Standardised	Interview	15	Scale	Gambling
Alcohol Dependence Data Questionnaire	Raistrick, 1983	Alcohol	Standardised	Self-report	15	Scale	Substance use
Alcohol Dependence Scale	Skinner, 1982	Alcohol	Standardised	Self-report	25	Nominal	Substance use
Depression Scale	Radloff, 1977	Other	Standardised	Self-report	20	Scale	Depression/anxiety/stress, family/social relationships, physical health
Christo Inventory for Substance Misuse Services	Christo, Spurrell & Alcorn, 2000	Drugs/alcohol	Standardised	Interview	10	Nominal	Crime, employment, general health, psychological health, psychosocial functioning, risk behaviours, substance use
Christo Inventory of Drugs	Christo, n.d.	Drugs/alcohol	Standardised	Interview	8	Nominal, scale	Substance use
Clinical Opiate Withdrawal Scale	Wesson & Ling, 2003	Drugs	Standardised	Interview	11	Nominal	Substance use
Drug Abuse Screening Test	Gavin, Ross, & Skinner,	Drugs/alcohol	Standardised	Self-report	20	Nominal	Crime, employment,

	1989						family/social relationships, psychological health
Drug Avoidance Self-Efficacy Scale	Martin, Wilkinson, & Poulos, 1995	Drugs/alcohol	Standardised	Self-report	16	Scale	Self-efficacy
European Addiction Severity Index (EuropASI)	Blacken et al., 1994	Drugs/alcohol	Standardised	Interview	189	Nominal, scale	Crime, employment, family/social relationships, money, psychological health, substance use
Evaluate Your Alcohol Consumption	Sobell & Sobell, 1992	Alcohol	Standardised	Self-report	1	Scale	Substance use
Evaluate Your Drug Consumption	Sobell et al., 1996	Drugs	Standardised	Self-report	1	Scale	Substance use
The Health Questionnaire	Brodman et al., 1949	Health	Standardised	Self-report	144	Nominal	Physical health
Leeds Dependence Questionnaire	Raistrick et al., 1994	Drugs/alcohol	Standardised	Self-report	10	Scale	Substance use
Lifetime Drinking History	Skinner, 1982	Alcohol	Standardised	Interview	8	Nominal, scale	Substance use
Maudsley Addiction Profile	Marsden et al., 1998	Drugs/alcohol	Standardised	Interview	61	Nominal, scale	Crime, employment, family/social relationships, physical health, psychological health, psychosocial functioning, risk behaviours, substance use
Measurements in the Addictions for Triage and Evaluation	Schippers, Broekman, Buchholz & Rutten, 2009	Drugs/alcohol	Standardised	Interview	114	Nominal, scale	Community/civic life, depression/anxiety/stress, domestic life, employment, family/social relationships, money, personality, physical health, school, self-care, social life,

							substance use
Monitoring Area and Phase System - Intake	Öberg et al., 1997	Drugs/alcohol	Individualised	Interview	7	Nominal	Crime, employment, family/social relationships, physical health, psychological health, substance use
Monitoring Area and Phase System - Out	Öberg et al., 1997	Drugs/alcohol	Standardised	Interview	223	Nominal, scale	Crime, employment, family/social relationships, physical health, psychological health, substance use
Monitoring Area and Phase System - Up	Öberg et al., 1997	Drugs/alcohol	Standardised	Interview	102	Nominal, scale	Crime, employment, family/social relationships, physical health, psychological health, substance use
Objective Opiate Withdrawal Scale	Handelsman et al., 1987	Drugs	Standardised	Observational	13	Nominal	Substance use
Opiate Treatment Index	Darke, Ward, Hall, Heather & Wodak, 1991	Drugs	Standardised	Interview	104	Nominal, scale	Crime, physical health, use
Personal Experience Screening Questionnaire	Winters, 1992	Drugs/alcohol	Standardised	Self-report	40	Scale	Psychosocial functioning, substance use
Rosenberg Self-Esteem Scale	Rosenberg, 1965	Other	Standardised	Self-report	10	Scale	Self-esteem
Severity of Dependence Scale	Gossop et al., 1995	Drugs	Standardised	Self-report	5	Scale	Substance use
Goals of Treatment Questionnaire	Joosten et al., 2009	Drugs/alcohol	Standardised	Mixed	24	Nominal	Child care, community/civic life, crime, daytime activities, domestic life, family/social relationships, food

							and nutrition, gambling, housing status, money, physical health, psychological health, school, self-control, sexual functioning, social life, substance use, technology and information
Situational Confidence Questionnaire	Annis & Martin, 1985	Alcohol	Standardised	Self-report	39	Scale	Emotions, employment, physical health, self-control, social life
Situational Confidence Questionnaire - Heroin	Barber, Cooper & Heather, 1991	Drugs	Standardised	Self-report	22	Scale	Emotions, family/social relationships, physical health, self-control, social life
Subjective Opioid Withdrawal Scale	Handelsman et al., 1987	Drugs	Standardised	Self-report	16	Scale	Substance use
The Texas Christian University - Initial Assessment - Methadone Outpatient	Texas Christian University, 1995	Drugs/alcohol	Standardised	Interview	58	Nominal, scale	Employment, psychosocial functioning, substance use
The Texas Christian University Intake - Methadone Outpatient	Texas Christian University, 1997	Drugs/alcohol	Standardised	Interview	119	Nominal, scale	Acculturation, crime, employment, family/social relationships, gambling, general health, money, psychological health, psychosocial functioning, risk behaviours, substance use
Texas Christian University Follow - Up Interview - Methadone Outpatient	Texas Christian University, 1997	Drugs/alcohol	Standardised	Interview	84	Nominal, scale	Crime, employment, family/social relationships, gambling, general health, leisure activities, money, psychological

							health, risk behaviours, substance use
The Texas Christian University - Initial Assessment-Correctional Residential	Texas Christian University, 1998	Drugs/alcohol	Standardised	Interview	114	Nominal, scale	Employment, psychological health, psychosocial functioning, substance use
The Texas Christian University - Intake - Correctional Residential	Texas Christian University, 1998	Drugs/alcohol	Standardised	Self-report	95	Scale	Psychological health, psychosocial functioning
The Texas Christian University HIV/AIDS Risk Assessment	Texas Christian University, 1997	Drugs/alcohol	Standardised	Interview	19	Scale	Risk behaviours
The Texas Christian University Self Rating at Intake - Methadone Outpatient	Texas Christian University, 1996	Drugs/alcohol	Standardised	Interview	95	Scale	Psychological health, psychosocial functioning
The Texas Christian University-Evaluation of self and treatment	Texas Christian University, 1996	Drugs/alcohol	Standardised	Self-report	126	Scale	Crime, employment, family/social relationships, housing status, risk behaviours, substance use
The WHO Quality of Life-Bref instrument	Skevington, Lotfy, & O'Connel, 2004	Other	Standardised	Self-report	26	Scale	Family/social relationships, psychological health, social life
Timeline Followback Method	Robinson, Sobell, Sobell, & Leo, 2014	Alcohol	Standardised	Self-report	1	Scale	Substance use

Table 3: *Number and percentage of outcome measures featuring the 54 patient-generated sub-theme of problems*

Patient-generated sub-themes	No. instruments with sub-theme present (%)
Addiction	28 (67)
Work-related problems	15 (36)
Relationships difficulties: family-general	13 (31)
Money worries	7 (17)
Relationship difficulties: family - worry about another	15 (36)
Justice-related problems	15 (36)
Worries about health	10 (24)
Self image/self worth	1 (2) *
Coping:daily living	7 (17)
Loneliness/being alone	2 (5) *
Global	0 (0) *
Depression/Anxiety	15 (36)
Motivation	4 (9) *
Emotions – unspecified	2 (5) *
Relationship difficulties: family - breaking up	15 (36)
Relationships-general	14 (33)
Relationship difficulties: family – conflict	1 (2) *
Socialising	8 (19)
Agression/irritability	1 (2) *
Housing worries	2 (5) *
Relationship difficulties partner - breaking up	15 (38)
Communication	5 (12)
Relationship difficulties: family – caring	1 (2) *
Being happy	2 (5) *
Dependence on other people	0 (0) *
Guilt	0 (0) *
Suicidal thoughts	9 (19)
Understanding self/events	0 (0) *
Existence/existential	0 (0) *
Future	0 (0) *
Relationship difficulties: partner – conflict	2 (5) *
Concentration	9 (21)
Coping: general	0 (0) *
Fears/panics	3 (7) *
Moving on	0 (0) *
OCD	2 (5) *

Personal development	0 (0) *
Sleep problems	2 (5) *
Studies-related problems	2 (5) *
Victim of abuse/sexual violence	1 (2) *
Achievement	1 (2) *
Attempted Suicide	4 (9) *
Bereavement	0 (0) *
Coping: feelings	0 (0) *
Eating problems	1 (2) *
Going out/travelling	6 (14)
Having positive Outlook	0 (0) *
Having time	0 (0) *
Outlook on life	0 (0) *
Relationship difficulties: partner – development	15 (36)
Relationship difficulties: partner – general	6 (14)
Relationship difficulties: partner - worry about another	15 (36)
Self-acceptance	1 (2) *
Sexual problems	1 (2) *

Note: The values highlighted with a * represent sub-themes covered by <10% of outcome measures.

Table 4*Dimensions (of patient-generated sub-themes) covered by the outcome measures, based on the multiple correspondence analysis*

Variables (patient-generated sub-themes)	Dimensions							
	1		2		3		4	
	Discrimination	Contribution (%)	Discrimination	Contribution (%)	Discrimination	Contribution (%)	Discrimination	Contribution (%)
Addiction	0.003	0.0	0.021	0.7	0.437	22.0	0.155	11.1
Communication	0.000	0.0	0.589	18.7	0.252	12.7	0.079	5.6
Concentration	0.214	2.5	0.359	11.4	0.252	12.7	0.024	1.7
Coping: daily living	0.002	0.0	0.397	12.6	0.180	9.1	0.001	0.1
Depression/Anxiety	0.459	5.4	0.239	7.6	0.085	4.3	0.000	0.0
Going out/travelling	0.012	0.1	0.607	19.3	0.233	11.7	0.023	1.6
Justice-related problems	0.498	5.9	0.002	0.1	0.001	0.1	0.236	16.8
Money worries	0.397	4.7	0.053	1.7	0.002	0.1	0.103	7.4
Relationship difficulties: family - breaking up	0.867	10.3	0.074	2.3	0.012	0.6	0.027	1.9
Relationships difficulties: family-general	0.531	6.3	0.008	0.3	0.003	0.2	0.032	2.3
Relationship difficulties: family - worry about another	0.867	10.3	0.074	2.3	0.012	0.6	0.027	1.9
Relationship difficulties: partner - breaking up	0.867	10.3	0.074	2.3	0.012	0.6	0.027	1.9
Relationship difficulties: partner – development	0.867	10.3	0.074	2.3	0.012	0.6	0.027	1.9
Relationship difficulties: partner – general	0.268	3.2	0.001	0.0	0.029	1.5	0.028	2.0
Relationship difficulties: partner - worry about another	0.867	10.3	0.074	2.3	0.012	0.6	0.027	1.9
Relationships-general	0.579	6.9	0.011	0.3	0.021	1.1	0.003	0.2

Socialising	0.108	1.3	0.086	2.7	0.083	4.2	0.373	26.6
Suicidal thoughts	0.187	2.2	0.400	12.7	0.188	9.5	0.095	6.8
Work-related problems	0.466	5.5	0.009	0.3	0.132	6.7	0.035	2.5
Worries about health	0.366	4.3	0.000	0.0	0.024	1.2	0.077	5.5
Active total	8.424	100.0	3.149	100.0	1.983	100.0	1.401	100.0
Inertia / Variance	0.421		0.158		0.099		0.070	

Note: The values, which vary between 0 and 1, indicate how much each variable contributes to/is present in each dimension (i.e., higher values indicate a greater contribution).

Figure Legends

Figure 1: *Flowchart showing the selection of outcome measures for review*

Figure 2: *List of domains ($N = 31$) ordered by frequency of outcome measures featuring each domain*

Figure 3: *A similarity tree grouping the 42 outcome measures, based on their proximity in terms of domains covered*

Figure 1.

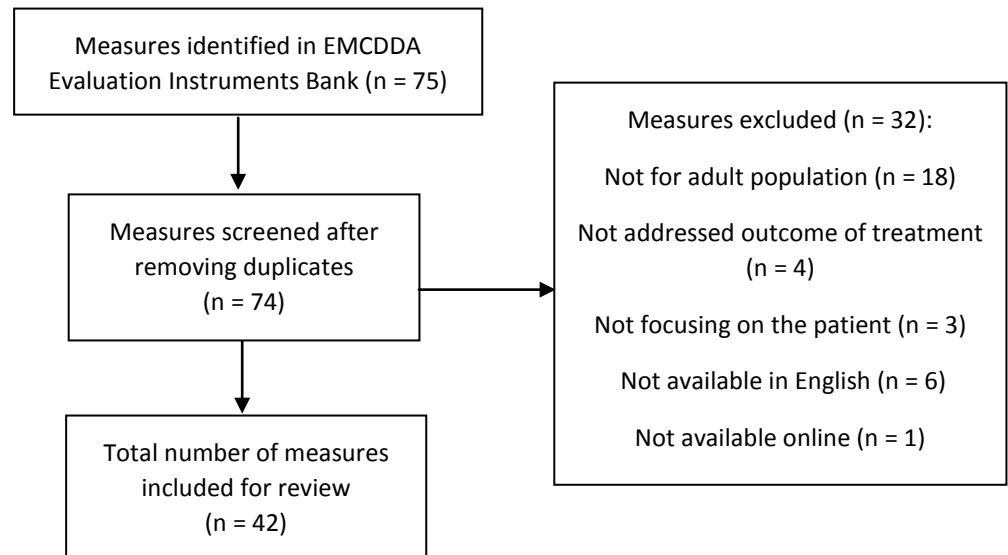


Figure 2

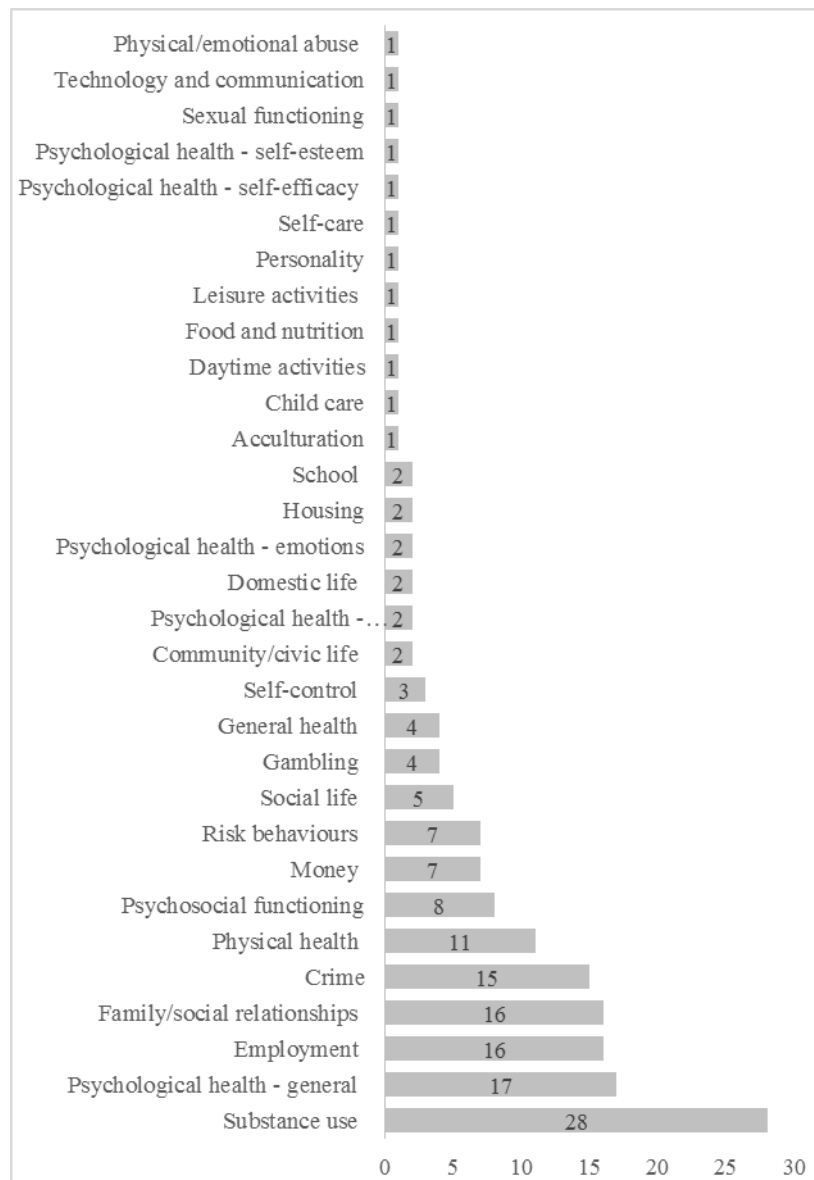


Figure 3

