
Article:
Barrowcliffe, A.L., Oathamshaw, S.C. and Evans, C. orcid.org/0000-0002-4197-4202 (2018) Psychometric properties of the clinical outcome routine evaluation-learning disabilities 30-item (CORE-LD30). Journal of Intellectual Disability Research, 62 (11). pp. 962-973. ISSN 0964-2633

https://doi.org/10.1111/jir.12551

ABSTRACT

Background There is paucity in availability of valid and reliable measures of psychopathology that can be routinely applied with an ID population in clinical practice. The psychometric properties of the Clinical Outcome Routine Evaluation-Learning Disabilities 30-Item version (CORE-LD30) are examined.

Method The CORE-LD30 was administered to 271 sequential referrals to three NHS ID services providing psychological support. A Principal Components Analysis with oblique rotation was conducted with examination of convergent validity for extracted domains.

Results Three rotated factors were extracted with good levels of internal consistency reported for the overall measure (α=.92) and each of the domains, conceptually labelled Problems/Symptoms (PS; α=.90), Risk to Self (RS; α=.76) and Risk to Others (RO; α=.71). Convergent validity is reported for two domains (PS with the GDS-LD, and RO with the HoNOS-LD), and support for the CORE-LD30 as a 'core' measure indicated.

Conclusion The CORE-LD30 is recommended as a useful broad ranging measure of psychopathology for use with an ID population. Domains may prove to be useful for research and clinical purposes. Further research is recommended to examine the ability to monitor clinical change associated with specific levels of presentation and different clinical presentations/cohorts.

Key words

Intellectual Disability; Learning Disability; outcomes; CORE-LD; assessment

INTRODUCTION

There is a wide variation in identified prevalence rates for mental health problems in people with intellectual disabilities, depending on the method used to establish diagnosis (e.g. case notes review, clinical assessment, structured diagnostic schedules), the diagnostic classification system used (e.g. ICD-10-DCR, DSM-IV-TR, DC-LD) and the population studied (e.g. in-patient or community, nationality of clinical cohorts) (Taylor et al. 2008). The different research methodologies used make it more difficult to establish whether people with intellectual disabilities experience mental health problems at a different rate than the general population, with the exception of psychosis where studies have generally found the prevalence rate to be higher (Cooper et al. 2007). However, people with intellectual disabilities are more likely to experience a range of psychosocial and economic circumstances strongly associated with a risk of developing mental health problems, including unemployment and poverty, lack of social support and an intimate relationship (Emerson et al. 2001) and as a result studies using screening assessments indicating equivalent or higher rates of common mental health problems in this population (Hatton & Taylor, 2005) are likely to be valid. A recent NICE Guideline (September 2016) suggests a lifetime prevalence rate of 28% for mental health problems in the population of people with intellectual disabilities when problem behaviours are excluded, which is significantly higher than for people who do not have intellectual disabilities and with psychosis, dementia, autism and attention deficit hyperactivity disorder all being more common than in the general population.

In response to the evidence that mental health problems are common in this population, the last 20 years has seen a growth of research supporting the use of psychological therapy for a range of mental health problems in people with

intellectual disabilities including depression, anxiety, psychosis and anger (McGillivray et al. 2007, Lindsay et al. 1989, Haddock et al. 2004, Willner et al. 2002). This research is of sufficient quality for psychological therapies to be recommended for treating these difficulties in people with intellectual disabilities as well as the general population (Scottish Government, 2011, 2015).

Psychological therapies and Intellectual Disabilities

Cognitive-behavioural therapy has become increasingly available in intellectual disability services since the late 1990's when one-third of British psychologists reported using it frequently (Nagel & Leiper, 1999) and evidence that people with intellectual disabilities can make some of the links between activating events, thoughts and feelings required to engage in CBT (Dagnan et al. 2000, Oathamshaw & Haddock, 2006) support this development. Other types of psychological therapy have been used with this population for a number of years including psychodynamic therapy (Beail & Jackson, 2013) and behavioural therapy (Jones & Dowey, 2013). In recent years mindfulness-based approaches (Singh et al. 2013) have become popular in intellectual disability services and unusually this development has been in parallel, instead of a number of decades behind, the development of mindfulness-based therapy in the general population.

Assessment measures and Intellectual Disabilities

The difficulties in detecting and diagnosing mental health problems in people with intellectual disabilities using standard diagnostic criteria, which is one reason for the widely varying prevalence rates reported, have led to the development of a range of assessment tools modified from those used with the general population such as the DC-LD (Royal College of Psychiatrists, 2001), SCL-90-R (Derogatis, 1983; Kellett,

1999) and BSI (Derogatis, 1993, Kellett et al. 2003, 2004), while others are population-specific such as the PIMRA (Matson et al. 1984), Reiss Screen (Reiss, 1988) and PAS-ADD schedules (Moss et al. 1997). Several of these assessments use an informant reflecting the difficulties some people with intellectual disabilities have in communicating their symptoms and reporting on internal states, though at least for people with mild learning disabilities there is research evidence that they can provide a good account of their symptoms (Dagnan & Lindsay, 2004).

As psychological therapies for people with intellectual disabilities have become more widely available in clinical practice there is an increasing need to develop valid and reliable assessments to evaluate the effectiveness of therapy and serve as outcome measures, both problem specific and trans-diagnostic (Lindsay et al. 2015). In the UK, commissioners in England are expecting evidence of better outcomes for people with intellectual disabilities (e.g. Royal College of Psychiatrists, 2012) and in Scotland the Scottish Government has introduced an 18 week referral-to-treatment target (Health Efficiency Access to Treatment, but more commonly known by the acronym HEAT) for psychological therapies and an expectation that therapy outcomes are reported and monitored (Scottish Government, 2011, 2015).

Problem-specific measures with good psychometric properties have been developed for depression (Glasgow Depression Scale, Cuthill et al. 2003), anxiety (Glasgow Anxiety Scale-ID, Mindham & Espie, 2003) and measures used with the general population for assessing symptoms of psychosis (Psychotic Symptom Rating Scales PSYRATS), Haddock et al. 1999; Positive and Negative Syndrome Scale (PANSS), Kay et al. 1989) have been validated for people with intellectual disabilities (Hatton et al. 2005). While problem-specific measures are useful there remains a need to develop trans-diagnostic and pan-theoretical outcome measures that can be used

across a range of mental health problems and therapies and meet the requirements specified by Thornicroft and Slade (2000); 1) standardised, 2) acceptable to clinicians and, 3) feasible for ongoing routine use.

Vlissides and colleagues (2017) recently developed the Psychological Therapies Outcome Scale – Intellectual Disabilities (PTOS-ID) as a trans-diagnostic and trans-theoretical outcome measure designed to be used in routine clinical practice. This novel measure has been subject to an assessment of construct validity applying a principal components analysis and following orthogonal rotation, three factors emerged (anger and mood; positive well-being; anxiety). Internal consistency and concurrent validity were also assessed through a comparison with the Brief Symptoms Inventory (BSI; Derogatis, 1993).

The Clinical Outcomes in Routine Evaluation (CORE) system

The Clinical Outcomes in Routine Evaluation-Outcome Measure (CORE-OM) (Evans et al. 2000) is a pan-theoretical measure of psychological distress that is widely used as a routine outcome measure in mental health services across the UK and has good psychometric properties (Evans et al. 2002). The CORE-OM contains items assessing wellbeing and functioning as well as mental health problems. The authors were careful to explain that these four domains were never expected to form a clean factor structure in cross-sectional data and their exploratory factor analysis found a large first factor of negative problem items accounting for 38% of the variance, and smaller factors for positive statements and risk items (Evans et al. 2002). An intellectual disability derivative of the CORE-OM (CORE-LD) was developed by two research groups using simplified questions and including novel items designed to assess experiences of having a learning disability, known as the 'missing domain'

(Brookes & Davies, 2007, Marshall & Willoughy-Booth, 2007). Using an intellectual disability participative research group and focus groups, the 34 CORE-OM items were reduced and modified and new items included to reflect a 'missing domain' relating to the lived experience of people with intellectual disabilities and their well-being (Brookes and Davies, 2007).The five point ordinal scale was reduced to three points, and symbols drawn from the Council of Edinburgh Bonnington Symbol System illustrating the questions were included.

Two versions of the LD-CORE emerged, one from the research group in Barnet of 14-items and a 30-item version developed by the research group in Scotland. Initial psychometric evaluation of the 14-item version found it to have good test-retest reliability (Brookes et al. 2013) and the 30-item version was found to positively correlate with the GDS and GAS-ID, and to again to have good test-retest reliability (Marshall et al. 2013). Both versions of the CORE-LD are used in services, although AB and SO, as clinicians working with clients with ID, prefer the longer version as the greater number of problem items in it seems more comprehensive for evaluating pan-theoretical therapies in diverse routine clinical practice settings and as it remains closer in structure and item inclusion to the CORE-OM used extensively in non-ID mental health services.

However, the factor structure of the CORE-LD30 has not been explored and the 30-item version continues to use domains in the CORE-OM plus the 'missing domain', but without empirical exploration of the domain structure of this significantly modified version. The principal aim of this study is to examine the factor structure of the CORE-LD30 and then to explore convergent and divergent validity against other measures.

METHODOLOGY

Contributing services

Clinical data was obtained from three specialist Adult Learning Disability Services within the NHS collaborating in an analysis of data collected as standard practice within each contributing team. Minimal demographic data were shared, with the main focus of the analysis being in regard to the structure of the CORE-LD30 in a naturalistic clinical setting. Data correspond to 271 sequential referrals into these services where the service user was able to complete a CORE-LD30 as part of the standard assessment protocol and where Intellectual Disability (ID) status had been confirmed. The study received approval from the local Audit Departments in each participating Trust and the National Caldicott Guardian Scrutiny Panel for data sharing.

Psychometric measures

A number of different psychometric assessment measures in addition to the CORE-LD30 were used across services as part of assessment protocols, dependent upon reasons for referral and presenting needs. Whilst this means that not all participants completed a full range of measures reported in the analysis, it did mean that a broader range of measures were available for the evaluation of concurrent validity based on data obtained from standard assessments in a naturalistic context.

Available measures consisted of the Health of the Nation Outcome Scale for Learning Disabilities (HoNOS-LD; Roy et al. 2002), the Glasgow Depression Scale for people with a Learning Disability (GDS-LD; Cuthill et al, 2003), the Glasgow Anxiety Scale for people with an Intellectual Disability (GAS-ID; Mindham & Espie, 2003) and the State-Trait Anger Expression Inventory (STAXI-2; Spielberger, 1999).

The HoNOS-LD (Roy et al. 2002) is comprised of 18-items and designed for use with people with a learning disability and mental health needs as a global outcome measure (e.g. Hillier et al. 2010). Although it is recognised that it may be insensitive to subtle changes in presentation for people with learning disabilities and chronic conditions, it is an established reliable and validated measure of overall functioning for this clinical cohort (Roy et al. 2002).

The STAXI-2 (Spielberger, 1999) is comprised of 57 questions across three sections, 'how I feel right now', 'how I generally feel', and 'how I generally react when angry or furious'. Only items from the latter 2 sections were included in the overall analysis, as these responses contribute to an overall anger expression score (the AX-Index). The AX-Index score provides an overall estimate of an individual's propensity to express anger either externally to others or inwardly to the self. Whilst the STAXI-2 has not been empirically validated for use with adults with ID, its use has been reported elsewhere with this population (e.g. Taylor et al. 2005). Additionally, it has been used with minor question adaptation in contributing services for many years (guidelines are available from the authors on request).

The GDS-LD (Cuthill et al. 2003) is a 20-item self-report measure of depression developed specifically for use with learning disability populations, whilst the GAS-ID (Mindham & Espie, 2003) comprises 27-items for the assessment of anxiety. Good levels of validity and reliability are reported for both instruments (Cuthill et al. 2003; Mindham & Espie, 2003).

RESULTS

Data were analysed using SPSS Version 24, with effect size calculations computed using http://socscistatitistics.com/effectsize/default3aspx.

Participants

Two-hundred and seventy-one people with an established intellectual disability had completed the CORE-LD30 as part of the standard assessment undertaken within participating teams. Each participating team was multidisciplinary in composition (e.g. National LD Professional Senate Guidance, 2015) although only those referred for psychological therapies were administered the CORE-LD30 where they were able to self-report symptoms during the course of clinical interview. The sample had an average age of 34.2 years (SD = 12.5) with a range of 17-70 years. It comprised 140 men (mean = 32.3 years; SD = 12.3) and 131 women (mean = 36.3 years; SD = 12.5), with a significant age difference by gender ($F(1, 269) = 7.2$, $p = 0.008$).

Suitability of data for factor (principal component) analysis

The inter-correlation of CORE-LD30 items was examined. Acceptable levels of inter-correlation between variables for factor analysis (Bartlett's Test $p < .001$) and the Kaiser-Myer-Olkin measure (KMO) of .90) indicate adequate item communality and sampling. The sample size of 271 was sufficient for extraction of a possible four factors based on simulation work (Guadagnoli & Velicer, 1988 and Velicer & Fava, 1998).

Principal components analysis and extraction of factors

A principal components analysis was performed utilising all 30 items. The point of inflection on the Scree Plot (Figure 1) suggested reduction to three components to rotate and that seemed congruent with analyses (noted above) of the CORE-OM. The oblimin rotation seemed to reveal three components with acceptable face validity based on the explained variances and consideration of items loading on each

component (e.g. Costello & Osborne, 2005; Field, 2005; Tinsley & Tinsley, 1987). These three components accounted for 41.4% of the overall explained variance.

Insert Figure 1 about here

The pattern matrix after rotation shows a separation of items into three domains, designated as: (1) Problems/Symptoms; (2) Risk to Self; (3) Risk to Others. Regression coefficients (> .35) are presented in the pattern matrix (Table 1) with respective items. Cross-loading is only observed for one item and considered at an acceptable level (Costello & Osborne, 2005). Examination of the Component Correlation Matrix (Table 2) supports oblique rotation, with moderate levels of relatedness observed between components. Good levels of internal reliability as measured by Cronbach's alpha are observed for each domain and the overall measure (Table 3).

Insert Tables 1 – 3 about here

Convergent validity

Pearson correlations between domains (scores based on means of item scores within the component) *versus* scores on the completed psychometric measures were calculated. As previously noted, only subsets of participants completed additional measures in addition to the CORE-LD30 during the course of initial assessments, reflecting a range of clinical presentations and needs addressed within Specialist Community Learning Disability teams and the fact that data is utilised from an opportunistic sample. The scoring of each of the measures described has similar directionality, in that higher scores indicate greater levels of reported symptom presentation. As observed in Table 4, all domains correlate significantly with the

HoNOS-LD, GDS-LD, and the AX-Index scores. Unsurprisingly, given the sample size of 28, the correlations with the GAS-ID are less clear. The cross-correlation, between the risk to self score of the CORE-LD30 and the GAS-ID fails to meet conventional statistical significance (lower 95% confidence limit = -.14) and the correlation of the risk to others score of the CORE-LD30 and the GAS-ID is negative at -.35. Clearly it would be very desirable to have larger samples in which the CORE-LD30 and GAS-ID were both used to have more precision on these correlations but it seems likely that the low correlations have arisen because the GAS-ID does not focus on risks to others. These correlations should be appraised with those between the other measures, shown in Table 5.

Insert Table 5 about here

Gender differences

As displayed in Table 6, results indicate a small gender difference for the Problems/ Symptoms domain and for the overall CORE-LD30 scores, with females scoring significantly higher than males in both cases. No statistically significant gender differences were observed for other measures employed (Table 7) but the broad 95% confidence intervals (CI's) resulting from relatively small samples completing these sets of additional measures should be noted.

Insert Tables 6 & 7 about here

DISCUSSION

The main purpose of this study was to examine the psychometric properties of the CORE-LD30 and relate findings to its utilisation as a measure in the spirit of "core: something onto which other measures can be added" (Evans et al., 2002, p.59). A

principal components analysis with oblique rotation resulted in the final extraction of three separate domains and an indication that the overall measure does indeed function appropriately as a 'core' measure. The emergent domains are interestingly different to those based on the derivation of the LD-CORE from the CORE-OM and some differences from the broad structure of the CORE-OM, "a complex factor structure and may be best scored as 2 scales for risk and psychological distress" (Lyne et al. 2006, p.185). The main difference is that in these CORE-LD30 data the risk domain of the CORE-OM seems to be clearly split across two domains: risk to self and risk to others.

Are the CORE-LD30 domains robust?

In regard to the convergent validity of extracted components, we observe 'medium' to 'strong' correlation coefficients when reviewing 95% CI's for two of our three component derived scores (Cohen, 1988) against other measures: specifically the Problems/Symptoms and the Risk to Others domains. For the Problems/Symptoms score we observe a significant correlation with acceptable 95% CI's with the GDS-LD, indicating that this domain is sensitive to depressive symptom presentation. Support for the meaningful independence of the Risk to Others domain is indicated in the strength of the correlation (within 95% CI's) with the HoNOS-LD scores. This suggests that the severity of self-reported mental health symptoms is positively correlated with items suggestive of elevated risk to others. Items within the Risk to Others domain indicate cognitive (e.g. perceptions of persecution), affective (e.g. negative mood states) and behavioural (e.g. verbal and/or physical attack) features associated with elevated risk to others. Clearly further examination of the Risk to Others scores against observed/reported behaviour would be invaluable in further considering the clinical validity of this domain.

Where we do observe domain specific correlations, we also observe similar positive correlations (within 95% CI's) for the overall CORE-LD30 values. Thus we observe positive correlations between the total CORE-LD30 score and the GDS-LD, and the total CORE-LD30 score and the HoNOS-LD. These demonstrate that the overall CORE-LD30 values are also sensitive to levels of self-reported depressive symptomatology and general mental health presentation. Support for the GDS-LD and HoNOS-LD measuring rather different things is given from their correlation (Table 5) of .14 (95% CI -.15 to .41). This supports the utility of the CORE-LD30 as a broad-ranging 'core' measure.

None of the psychometric measures available for statistical analysis with this routine clinical sample appeared to be strongly correlated with the Risk to Self domain when looking at the lower limits of the 95% CI's. This doesn't necessarily undermine the clinical utility of this domain as an independent feature of the CORE-LD30, but suggests further examination with more targeted measures of hopelessness and suicidality may be required to assess convergent validity. The fact that there are independent scales for depression and hopelessness available for use with the general population (e.g. Beck, 1988; Beck, Steer & Brown, 1996) indicates that general symptoms of low mood and hopelessness are recognised as independent of one another. Further assessment of this would be valuable in future studies, to establish if there are similar findings in a clinical population of people with intellectual disability.

Previous studies (e.g. Marshall, Coiffait & Willoughby-Booth, 2013) have reported statistically significant correlations between the CORE-LD30 and the GDS-LD and the GAS-LD, concluding that the CORE-LD30 was appropriate for assessing levels of depression and anxiety, although no confidence intervals were reported. As our

findings show, correlation coefficients reported in the absence of confidence intervals for relatively small samples, risk over-confident interpretation of correlations.

Gender differences

We observe small but significant gender effects for Problems/Symptoms and for overall CORE-LD30 scores, with higher scores observed for the female participants. Such directional gender differences in a clinical sample were reported on CORE-OM domains of well-being and risk only (Evans et al., 2002). We observe no such gender effect for other measures employed within this study, although note that this may be attributable to sample size (reflected in the large 95% CI's), but does indicate that the CORE-OM in original and adapted formats remains sensitive to gender.

Future research

A number of additional studies are recommended to further examine the value of this measure in clinical practice. This includes an examination of the sensitivity of the CORE-LD30 to change following delivery of clinical interventions, examination of clinical cut-offs indicating clinical severity across domains, and the concurrent validity of risk domains across different clinical diagnoses and cohorts. The internal validity of domains as demonstrated here supports use of the CORE-LD30 in routine clinical practice as a 'core' measure of clinical progress with valuable domain specific information now available. In particular, the relevance of domain scores to identified risk factors utilised in the stratification of 'admission avoidance' registers (NHSE, 2017) would be very interesting to explore.

Summary and conclusions

Significant differences on measures of convergent validity are supportive of the principal component based domains reported here. Whilst acknowledging that clear convergent validity is observed only for two of the three domains (Problems/Symptoms and Risk to Others), this is considered to be a limitation of the measures utilised in the current study. This may be indicative of a domain (Risk to Self) within the CORE-LD30 not identified in these other assessment/outcome measures used as standard practice in the services contributing data. Further studies designed to explicitly look at issues of concurrent validity for the identified domains would be recommended, noting that whilst we had access to a relatively large opportunity sample as required for the principle components analysis, only certain additional measures were clinically applicable to administer alongside the CORE-LD30.

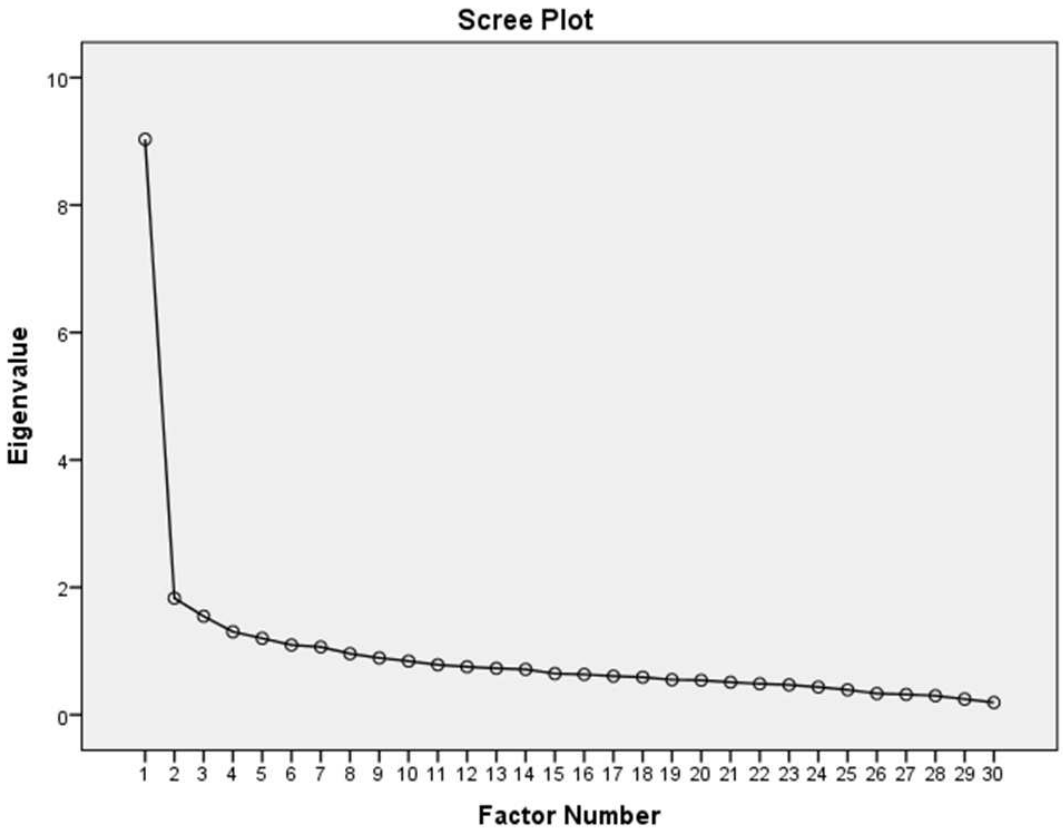

Figure 1. Scree plot

Table 1. Component loadings for the CORE-LD30 items

No	Item	1	2	3
1	Have you felt very, very lonely?	.50		
2	Have you felt really worried?	.58		
3	Have you felt confused?	.50		
4	Have you felt like you have no energy to do anything?	.44		
7	Have you felt frustrated on upset with your learning disability?	.58		
9	Have you found it hard to say how you feel?	.57		
10	Have you been too worried or anxious to do important things?	.63		
12	Have you felt sad about people you have lost?	.64		
13	Have you felt like crying?	.60		
14	Have you felt really scared and frightened?	.56		
15	Have you bottled up angry feelings?	.56		
16	Have your problems felt too much for you?	.61		
17	Have you had difficulty getting to sleep or falling asleep?	.42		
18	Did you think about your problems all the time?	.65		
23	Has it been hard to make friends?	.42		
24	Have you felt unhappy?	.55		
25	Have you been really upset by memories or pictures that pop into your head?	.70		
27	Have you thought that your problems were your fault?	.44		
6	Have you managed to cope when things went wrong?		.50	
8	Have you thought about hurting yourself?		.83	
11	Have you felt happy with things you have done?		.35	
20	Have you felt that life is hopeless?	.44	.50	
21	Have you thought about ending your life?		.82	
28	Have you thought things will get better?		.37	
30	Have you hurt yourself on purpose?		.64	
5	Have you attacked someone?			.73
19	Have you threatened or shouted at someone?			.69
22	Have you felt people are getting at you?			.49
26	Have you been cross or grumpy with other people?			.51
29	Have other people made you feel really bad about yourself?			.53

Table 2. Component Correlation Matrix

Component	1	2	3
1	1.000	.367	.329
2	.367	1.000	.280
3	.329	.280	1.000

Table 3. Cronbach's coefficient α and internal reliability of domain items

Domain	Number of items	Coefficient α
Problems/symptoms	18	.90
Risk to self	7	.76
Risk to others	5	.71
Total CORE-LD30 score	30	.92

Table 4. Pearson correlations between domains and other psychometric measures (95% Confidence Intervals)

	Problems/ symptoms	Risk to self	Risk to others	Total CORE-LD30 score
HoNOS-LD (*n*=107)	*r* = .39 (.22 to .54)	*r* = .32 (.14 to .48)	*r* = .51 (.35 to .64)	*r* = .47 (.30 to .60)
GDS-LD (*n*=94)	*r* = .67 (.54 to .77)	*r* = .43 (.25 to .58)	*r* = .43 (.25 to .58)	*r* = .69 (.56 to .78)
GAS-ID (*n*=28)	*r* = .59 (.27 to .79)	*r* = .25 (-.14 to .59)	*r* = -.35 (-.64 to .03)	*r* = .48 (.12 to .72)
AX-Index (*n*=41)	*r* =.52 (.25 to .71)	*r* = .31 (.01 to .57)	*r* = .34 (.03 to .58)	*r* = .54 (.28 to .73)

Table 5. Pearson correlation coefficients between other psychometric measures (95% Confidence Intervals)

	GAS-ID[1]	AX-Index	HoNOS-LD
GDS-LD	*r* = .44 (.05 to .72) *n*=24	*r* = .46 (.18 to .67) *n*=41	*r* = .14 (-.15 to .41) *n*=48
AX-Index			*r* = .12 (-.24 to .45) *n*=31

[1] Note that Pearson values for the GAS-ID could not be calculated for AX-Index and HoNOS-LD as *n*<2 in both cases

Table 6. Means and standard deviations for males and females across domains

Domain	Male (*n*=140)		Female (*n*=131)		95% CI	
	Mean	(s.d.)	Mean	(s.d.)	Difference	Cohen's d †
Problems/Symptoms	14.66	(8.47)	18.61	(8.24)	1.95 to 5.95	.12 to .36
Risk to Self	2.20	(2.28)	2.60	(2.49)	-0.17 to 0.97	-.04 to .20
Risk to Others	2.56	(2.20)	3.45	(2.40)	0.34 to 1.44	.07 to .31
All items	20.68	(12.03)	26.04	(11.78)	2.51 to 8.21	.10 to .35

† Interpretation of Cohen's d: < .1 'trivial'; .1 to .3 'small'; .3 to .5 'moderate'; >.5 'large'

Table 7. Means and standard deviations for males and females across measures

Measure	Male		Female		95% CI	
	Mean	**(s.d.)**	**Mean**	**(s.d.)**	**Difference**	**Cohen's d** †
HoNOS-LD	*n*=66		*n*=41			
	14.44	(7.89)	14.61	(7.86)	-2.94 to 3.28	-.18 to 0.20
GDS-LD	*n*=58		*n*=36			
	13.69	(7.27)	16.53	(7.11)	-0.20 to 5.88	-.01 to 0.40
GAS-ID	*n*=18		*n*=10			
	25.22	(11.69)	26.40	(9.95)	-7.83 to 10.19	-.32 to .43
AX-Index	*n*=28		*n*=13			
	51.89	(13.44)	52.31	(12.31)	-8.48 to 9.32	-.30 to .32

† Interpretation of Cohen's d: < .1 'trivial'; .1 to .3 'small'; .3 to .5 'moderate'; >.5 'large'

References

Barrowcliff A.L. & Haddock G. (2010) Factors affecting compliance and resistance to auditory command hallucinations: perceptions of a clinical population. *Journal of Mental Health* **19**, 542–52.

Beail N. & Jackson T. (2013) Psychodynamic Psychotherapy and People with Intellectual Disabilities. In: *Psychological Therapies for Adults with Intellectual Disabilities (*eds J.L. Taylor, W.R. Lindsay, R.P. Hastings & C. Hatton), pp 237-52. Wiley-Blackwell, Chichester, West Sussex.

Beck A.T. (1988) Beck Hopelessness Scale. The Psychological Corporation, San Antonio, TX

Beck A.T., Steer R.A. & Brown G.K. (1996) Manual for the Beck Depression Inventory-II. Psychological Corporation, San Antonio, TX

Brookes M. & Davis S. (2007) Pathways to participatory research in developing a tool to measure feelings. *British Journal of Learning Disabilities* **36**, 128-33.

Brookes M., Davis S. & Twigg E. (2013) A measure for feelings – using inclusive research to develop a tool for evaluating psychological therapy (Clinical Outcomes in Routine Evaluation – Learning Disability). *British Journal of Learning Disabilities* **41**, 320-29.

Cohen J. (1988) *Statistical power analysis the behavioral sciences* (2nd edition). Academic Press, New York.

Cooper S.A., Smiley E., Morrison, J., Williamson A & Allan L. (2007) Mental ill-health in adults with intellectual disabilities: prevalence and associated factors. *The British Journal of Psychiatry* **190 (1),** 27-35.

Costello A.B. & Osborne J. (2005) Best practice in exploratory factor analysis: four recommendations for getting the most from your analysis. *Practical Assessment, Research & Evaluation* **10(7)**. Available online at: http://pareonline.net/getvn.asp?v=10&n=7

Cuthill F.M., Espie C.A. & Cooper S.A. (2003) Development and psychometric properties of the Glasgow Depression Scale for people with a Learning Disability. *British Journal of Psychiatry* **182**, 347-53.

Dagnan D., Chadwick, P. & Proudlove J. (2000) Towards an assessment of suitability of people with mental retardation for cognitive therapy. *Cognitive Therapy and Research* **24**, 627-36.

Dagnan D. & Lindsay W.R. (1997) Cognitive behaviour therapy for people with learning disabilities: assessment and intervention. In: *Cognitive Behaviour Therapy for People with Learning Disabilities* (eds B. Kroese, D. Dagnan & Loumidis K), pp, 110-23, Routledge, London.

Derogatis L.R. (1983) SCL-90-R: Administration, scoring, and procedures manual II. Baltimore, Clinical Psychometric Research.

Derogatis L.R. (1993) BSI Brief Symptom Inventory: Administration, Scoring, and Procedure Manual (4th Ed), Minneapolis, M.N.

Emerson E. & Jahoda A.J. (2013) Social and Psychological Factors as Determinants of Emotional and Behavioral Difficulties. In: *Psychological Therapies for Adults with Intellectual Disabilities (*eds J.L. Taylor, W.R. Lindsay, R.P. Hastings & C. Hatton), pp 15-30. Wiley-Blackwell, Chichester, West Sussex.

Evans C., Mellor-Clark J., Margison F., Barkham M., Audin K., Connell J. et al (2000) CORE: Clinical Outcomes in Routine Evaluation. *Journal of Mental Health* **9(3)**, 247-55.

Evans C., Connell J., Barkham M., Margison F., McGrath G., Mellor-Clark J. et al. (2002) Towards a standardised brief outcome measure: psychometric properties and utility of the CORE-OM. *British Journal of Psychiatry* **180**, 51-60.

Field A. (2005) *Discovering statistics using SPSS*. Sage Publications, London.

Gilbert P. (1992) *Depression: The Evolution of Powerlessness*. Lawrence Earlbaum Associates: Hove; Guilford: New York.

Guadagnoli, E., & Velicer, W. F. (1988). Relation of sample size to the stability of component patterns. *Psychological Bulletin* **103**, 265–275.

Haddock G., McCarron, J., Tarrier N & Faragher E.B. (1999) Scales to measure dimensions of hallucinations and delusions: the psychotic symptom rating scale (PSYRATS). *Psychological Medicine* **29(4)**, 879-89.

Haddock G, Lobban F., Hatton C & Carson R. (2004) Cognitive-behaviour therapy for people with psychosis and mild intellectual disabilities: a case series. *Clinical Psychology and Psychotherapy* **11**, 282-98.

Hatton C. & Taylor J.L. (2005) Promoting healthy lifestyles: mental health. In: *Learning disability: A life-cycle approach to valuing people* (eds G. Grant, P. Goward, M. Richardson & P. Ramcharan), pp. 559-603. Press/McGraw Hill Educational Maidenhead, Open University.

Hatton C. Taylor J.L., Coldwell J, Crossley R. & Peckham N. (2005) The reliability and validity of general psychotic rating scales with people with mild and moderate intellectual disabilities: an empirical investigation. *Journal of Intellectual Disability Research* **49**, 490-50.

Hillier B., Wright L., Strydom A. & Hassiotis A. (2010). Use of the HoNOS-LD in identifying domains of change. *The Psychiatrist* **34**, 322-26.

Hurley A. (2008) Depression in adults with intellectual disability: symptoms and challenging behaviour. Journal Intellectual Disability Research **52 (11)**, 905-16.

Ingham B., Painter J., Roy A., Hastings R., Robinson S. & Green C. (2013) Development of a needs based payment system for specialist Learning Disability health services: Results of a pilot project. *Report for the Care Pathways and Packages Project*.

Jones S.P. & Dowey A. (2013) Behavioral Approaches to Working with Mental Health Problems. In: *Psychological Therapies for Adults with Intellectual Disabilities (*eds J.L. Taylor, W.R. Lindsay, R.P. Hastings & C. Hatton), pp 223-36. Wiley-Blackwell, Chichester, West Sussex.

Kay S.R., Opler L.A. & Lindmayer J.P. (1989) The Positive and Negative Syndrome Scale (PANSS): Rationale and standardisation. *The British Journal of Psychiatry* **155 (Suppl 7)**, 59-65.

Kellett S.C., Beail, N., Newman D.W. & Mosley E. (1999) Indexing psychological distress in people with intellectual disability: use of the symptom checklist-90-R. *Journal of Applied Research in Intellectual Disabilities* **12**, 323-34.

Kellet S.C., Beail, N., Newman D.W. & Frankish P. (2003) Utility of the Brief Symptom Inventory (BSI) in the assessment of psychological distress. *Journal of Applied Research in Intellectual Disabilities* **16**, 127-35.

Kellett S.C., Beail, N., Newman D.W. & Hawes A. (2004) The factor structure of the Brief Symptom Inventory: Intellectual disability evidence. *Clinical Psychology and Psychotherapy* **11**, 275-81.

Lindsay W.R., Baty F.J., Michie A.M. & Richardson I. (1989) A comparison of anxiety treatments with adults who have moderate and severe mental retardation. *Research on Developmental Disabilities* **10**, 129-40.

Lindsay W.R., Tinsley S., Beail N., Hastings R.P. Jahoda A., Taylor J.L. & Hatton C. (2015) A preliminary controlled trial of a trans-diagnostic programme for cognitive-behavioural therapy with adults with intellectual disability. *Journal of Intellectual Disability Research* **59(4)**, 360-9.

Lyne K., Barrett P., Evans C. & Barkham M. (2006) Dimensions of variation on the CORE-OM. *British Journal of Clinical Psychology* **45**, 185-203.

Marshall K. & Willoughby-Booth S. (2007) Modifying the clinical outcomes in routine evaluation measure for people who have a learning disability. *British Journal of Learning Disabilities* **35**, 107-12.

Marshall K., Coffait F-M. & Willoughby-Booth S. (2013) Assessing distress among people with intellectual disabilities. *Learning Disability Practice* **16(3)** 26-30.

Matson J.L., Kazdin A.E. & Senatore V. (1984) Psychometric properties of the psychopathology instrument for mentally retarded adults. *Applied Research in Mental Retardation* **5(1),** 81-9.

Mcgillivray J., McCabe M.P. & Kershaw M.M. (2007) Depression in people with intellectual disability: An evaluation of a staff-administered treatment program. *Research in developmental disabilities* **29,** 524-36.

Mindham J. & Espie C.A. (2003) Glasgow Anxiety Scale for people with an Intellectual Disability (GAS-ID): development and psychometric properties of a new measure for use with people with mild intellectual disability. *Journal of Intellectual Disability Research* **47(1)**, 22-30.

Moss S.C., Ibbotson B., Prosser, H. Goldberg D.P., Patel P. & Simpson N. (1997) Validity of the PAS-ADD for detecting psychiatric symptoms in adults with learning disability. *Social Psychiatry and Psychiatric Epidemiology* **32**, 344-54.

Nagel B. & Leiper R. (1999) A national survey of psychotherapy with people with learning disabilities. *Clinical Psychology Forum* **129**, 14–18.

National LD Professional Senate Guidance (2015) Delivering effective specialist community learning disability health team support. http://www1.bps.org.uk>files>public

NHS Education for Scotland. (2011) Mental Health in Scotland. A Guide to delivering evidence-based Psychological Therapies in Scotland. The Matrix. Accessed at: http://www.nes.scot.nhs.uk/media/20137/Psychology%20Matrix%202013.pdf

NHS Education for Scotland. (2015) The Matrix (2015) A Guide to Delivering Evidence-Based Psychological Therapies in Scotland. Accessed at: http://www.nes.scot.nhs.uk/education-and-training/by-discipline/psychology/the-matrix-(2015)-a-guide-to-delivering-evidence-based-psychological-therapies-in-scotland.aspx

NHS England (2017). Care and Treatment Reviews (CTRs): Policy and Guidance Including policy and guidance on Care, Education and Treatment Reviews (CETRs)

for children and young people. Accessed at: https://www.england.nhs.uk/wp-content/uploads/2017/03/ctr-policy-v2.pdf

NICE guideline (NG54) (September 2016) Mental health problems in people with learning disabilities: prevention, assessment and management. Accessed at: https://www.nice.org.uk/guidance/ng54

Reiss S.A. (1988) Reiss Screen for Maladaptive Behaviour. IDS Publishing Company, Worthington, Ohio.

Roy A., Matthews H., Clifford P., Fowler V. & Martin D. (2002) Health of the Nation Outcome Scales for People with Learning Disabilities (HoNOS-LD). *British Journal of Psychiatry* **180**, 61-66.

Royal College of Psychiatrists (2001) OP48. Diagnostic Criteria for Psychiatric Disorders for Use with Adults with Learning Disabilities/Mental Retardation.

Royal College of Psychiatrists (2012) Improving the health and wellbeing of people with learning disabilities: An evidence-based commissioning guide for Clinical Commissioning Groups (CCGs). Accessed at: http://www.rcpsych.ac.uk/pdf/RCGP%20LD%20Commissioning%20Guide%202012%2010%2009%20FINAL.pdf

Singh N.N., Lancioni G.E., Winton A.S.W., Singh A.S.A., Singh A.N.A. & Singh J. (2013) Mindfulness-Based Approaches. In: *Psychological Therapies for Adults with Intellectual Disabilities* (eds J.L. Taylor, W.R. Lindsay, R.P. Hastings & C. Hatton), pp 267-76. Wiley-Blackwell, Chichester, West Sussex.

Taylor J.L., Novaco R.W., Gillmer B.T., Robertson A. & Thorne I. (2005) Individual cognitive-behavioural anger treatment for people with mild-borderline intellectual

disabilities and histories of aggression: A controlled trial. *British Journal of Clinical Psychology* **44(3)**, 367–82.

Taylor J.L., Lindsay W.R. & Willner P. (2008) CBT for people with intellectual disabilities : emerging evidence, cognitive ability and IQ effects. *Behavioural and Cognitive Psychotherapy* **36(6)**, 723-33.

Thornicroft G. & Slade M. (2000) Are routine outcome measures feasible in mental health? *Quality in Health Care* **9**, 84.

Tinsley H. & Tinsley D. (1987) *Uses of factor analysis in counselling psychology research* **34(4)**, 414-424.

Velicer, W. F., & Fava, J. L. (1998). The effects of variable and subject sampling on factor pattern recovery. *Psychological Methods* **3**, 231–251.

Vlissides N., Beail N., Jackson T., Williams K. & Golding L. (2017) Development and psychometric properties of the Psychological Outcome Scale – Intellectual Disabilities (PTOS-ID). *Journal of Intellectual Disability Research* **616)**, 549-59.

Willner P., Jones J., Tamsy R. & Green G. (2002) A Randomized Controlled Trial of the Efficacy of a Cognitive-Behavioural Anger Management Group for Clients with Learning Disabilities. *Journal of Applied Research in Intellectual Disabilities* **15**, 224-35**.**