

Name of Tool	Level of Service/Case Management Inventory (LS/CMI)
Category	General Risk Assessment (Validated)
Author / Publisher	Andrews, Bonta and Wormith
Year	2004

Description

- The LS/CMI is a measure of risk and need factors with a case management component.
- It includes five assessment sections, three summary sections and three case management sections. Section 1 (General Risk/Needs Section) consists of 43 items that are grouped into 8 subsections (the Central 8). Other assessment sections include: Specific Risk/Need Factors, Prison Experience: Institutional Factors, Other Client Issues: Social, Health and Mental Health, and Special Responsivity Considerations. Structural issues like poverty, race, gender and age are addressed in the 'Responsivity Considerations' section ([Wormith and Bonta, 2018](#)).
- Risk is categorised into five levels; 'very low', 'low', 'moderate', 'high' and 'very high.'
- It is designed to assist professionals in management and treatment planning in justice, forensic, correctional, prevention, and related agencies.
- The LS/CMI is normed on Canadian and North American probation and institutional populations for males and females. Supplementary norms provided for use in Singapore and the UK.
- There is a feature in the instrument allowing the assessor to override the initial risk level ([Guay and Parent, 2017](#)).

Age Appropriateness

16+

Assessor Qualifications

Assessors must possess advanced training, certification and experience in psychological assessment or a related discipline, or satisfactorily complete a training course certified by the publishers. Can be used by a large range of professionals including social work and probation services.

Strengths

- Combines risk assessment and case management in a single assessment tool.
- It expands the traditional risk/need assessment instrument to a more comprehensive assessment by including non-criminogenic needs, prison experience and responsivity considerations.
- Assessors are able to identify strengths in the individual and his/her circumstances.
- It allows for a professional override of risk level based on an assessment of strengths and specific risk factors that are not captured in Section 1.

- The LS/CMI utilised normative data from circa 20,000 females in inmate and community settings in four countries in order to address gender-specific risk and responsivity issues., including mothering concerns, adult victimisation issues and protective strengths.

Empirical Grounding



- Developed in part from the LSI-R, a well validated tool with the developments informed by further research and consultation with practitioners. The LS instruments are grounded in the ‘General Personality and Cognitive Social Learning’ theory, maintaining that behaviour such as criminal conduct is learned from interactions with others. It is a general theory, advancing the argument that the Central Eight risk/need features apply across gender, age and race. (Andrews, Bonta and Wormith, 2004; [Wormith and Bonta, 2018](#)).
- Section 1 was informed by a re-analysis of LSI-R item data following research in Ontario, Canada that became called the LSI-Ontario Version. When the LSI-R was developed, GPCSL theory had not fully matured. For instance, the Emotional/Personal subcomponents of the LSI-R was overly concerned with general feelings of emotional distress and psychotic illness and underemphasised antisocial personality features. Grounding in the ‘Central Eight’ and the inclusion of a case management system formed the basis for the development of the LS/CMI ([Wormith and Bonta, 2018](#)).

Inter-Rater Reliability



a) UK Research

None at present.

b) International Research

- [Rettinger and Andrews \(2010\)](#) – the LS/CMI attained moderate to high inter-rater reliability estimates ranging from .65 for ‘financial problems’ to .91 for the composite general risk/need score in a sample of females.
- [Labrecque et al. \(2017\)](#) found that 21 items demonstrated good or strong levels of inter-rater agreement with an ICC exceeding .60. Seven items, however, gave an inadequate level of consistency, with an ICC of below .50. Three of these items are in the ‘Companions’ section; the remaining four items are across the Antisocial pattern, procriminal attitudes/orientation, leisure and family/martial sections. It is suggested that training around these items should be strengthened to address this.

Validation History

General Predictive Accuracy



a) UK Research

None at present.

b) International Research

- In order to test the tool in Pakistan, [Bhutta and Wormith \(2016\)](#) translated the LS/CMI into the national language, Urdu, and adjusted some of the items (e.g. education subscale modified to align with the lower literacy rate than that of Western countries). The revised tool was administered to 55 adults released on probation and found to be a useful indicator of recidivism. The only notable difference was that female probationers tended to score higher on the general risk/needs score, which the authors suggested could be a reflection of a cultural or gender bias.

- [Gordon, Kelty and Julian \(2015\)](#) tested the LS/CMI on individuals who offended in Australia and found that its total scoring yielded a significant although weak predictive utility with an AUC of .62. On this basis, the authors caution that the LS/CMI may not be the most suitable tool for measuring risk in Australia.

- The LS/CMI was adapted into a French version by [Guay \(2016\)](#). This version was utilised in a study by [Guay and Parent \(2017\)](#) to assess 3646 individuals with a sentence of less than two years. Good predictive accuracy was shown for new arrests and new convictions (AUC range of .70-.77 and .72-.77 respectively), with the exception of new convictions for other crimes (AUC equal to .66).

- [Wormith, Hogg and Guzzo \(2012\)](#) - moderate to large correlations observed between the General Risk/Need Score and general recidivism (.47), violent recidivism (.28) and sexual recidivism (.17).

- [Wormith et al. \(2007\)](#) - in a 10-year follow-up sample of 61 adult males, LS/CMI attained moderate accuracy (AUC) for the prediction of recidivism (any new conviction (.65), non-violent conviction (.62), violent conviction (.68) and any re-incarceration (.69). This tool was unable to predict sexual convictions (.49).

- A study by [Spence and Haas \(2015\)](#) in West Virginia found risk scores are strongly predictive of recidivism, even when controlling for other factors such as age, gender, and ethnicity.

- On average, logistic regression models predict that each 1 point increase in total risk score increases the odds of arrest by about 3-4% and the odds of incarceration by about 5-9%.

- LS/CMI risk scores are the strongest predictor of recidivism for both Day Report Centre clients and Department Of Corrections inmates. Using only LS/CMI

risk scores, it is possible to correctly predict recidivism in 60-70% of cases. The inclusion of other variables (i.e., age, race, etc.) increases predictive accuracy of recidivism.

- [Caldwell et al. \(2018\)](#) evaluated predictive validity of 19,344 probationer records in Nebraska over a five and a half year period. It was found that the LS/CMI predicted outcomes better for minorities (those who did not identify as White Europeans or were White but of Hispanic descent) than non-minorities (White European Americans or those of non-Hispanic descent). An experiment was conducted to explore whether officers showed prejudice in their scoring; findings showed total risk scores remained stable across different racial groups.

- [Olver and Kingston \(2019\)](#) found that the LS/CMI (called the LSI-Ontario Version in this paper) predicted violent and general recidivism in the overall sample and among specific diagnostic groups (schizophrenia, anxiety disorders and mood disorders). Predictive accuracy for violence specifically was smaller although still significant, suggesting the need to use a violence-specific tool in conjunction with this one.

Validation History

Applicability: Females



a) UK Research

None at present.

b) International Research

- [Dyck and colleagues \(2018\)](#) looked at the predictive validity of the LS/CMI with 136 Atlantic Canadian individuals. Whilst it was found to be a strong predictor of recidivism for males (AUC=.75), it was even better for females (AUC=.94) over an average of 3.42 year follow-up period.

- [Andrews et al. \(2012\)](#) - the LS/CMI composite score had a mean AUC of .83 for recidivism across five different samples of females. With gender found to have a significant effect on the validity of substance abuse, this was controlled for in analyses, resulting in an AUC of .79.

- [Rettinger and Andrews \(2010\)](#) - in a 57-month follow-up, the LS/CMI was able to discriminate between different risk categories for females within prison and community settings. Of the 411 women in the study, the higher risk females were responsible for 74% of all new offences. The

	LS/CMI general risk/need component generated AUC scores for both general and violent recidivism of .87 and .86 respectively.
--	--

Validation History	
Applicability: Ethnic Minorities	
a) UK Research	None at present.
b) International Research	<ul style="list-style-type: none"> • In a meta-analysis involving samples of Aboriginal individuals, the central eight risk factors attained small to moderate mean effect sizes in relation to general recidivism ranging from .19 (family/marital) to .56 (criminal history). Smaller mean effect sizes were observed for violent recidivism (Gutierrez et al., 2013). • Wormith, Hogg and Guzzo (2015) applied the LS/CMI to 9692 Aboriginal and 24, 758 non-Aboriginal individuals. Predictive accuracy was demonstrated for both sets, with AUC scores in the range of .64 for Aboriginals and .74 for the remainder. The authors advise that assessors should consider special circumstances when carrying out interviews with Aboriginal individuals, such as cultural heritage, jargon and dialect, communication styles and relational expectations.

Validation History	
Applicability: Mental Disorders	
a) UK Research	None at present.
b) International Research	<ul style="list-style-type: none"> • Andrews (1995) found that those diagnosed with mental disorders categorised as 'high / very high risk' on the LS/CMI had recidivism rates of up to 73% for any re-offending compared to 17% who were rated as 'very low' risk. • Within a larger study, Girard and Wormith (2004) included a sub-sample of 169 prison inmates with mental health problems (depression, psychosis, previous suicide threats/attempts). This sub-sample was found to score significantly higher on the General Risk/Need total score than those without mental health issues (mean effect sizes of 21.95 and 19.48 respectively).

Contribution to Risk Practice

- The developers of the tool indicated that the LS/CMI provides a 'gender-informed' assessment for risk, needs and responsivity issues; thus, it can be used across various settings without the need for separate gender-specific assessments.
- The LS/CMI aids the assessor in identifying risk, need and responsivity factors relevant to the individual's likelihood of re-offending and of other issues relevant to a holistic case management plan.
- Many of the factors identified within the assessment can act as targets for treatment/change and the tool can aid assessors in determining the level of monitoring and supervision required with regards to the formulation of case management plans.
- LS/CMI has an ability to highlight the strengths of the individual. These are factors that would actively enable the individual to desist from further offending and enables assessors to provide further information on these strengths in relation to potential clinical override of the level of risk generated from Section 1.
- Assessors are given the chance to elaborate on factors which have been highlighted as a strength in the 'General Risk/Needs' section (Andrews, et al. 2004: 153).
- Literature also describes how the LS/CMI may be used in recommendations for sentencing (see [Wolbransky et al., 2012](#)). A study looking at the correspondence between presentence risk evaluations and sentencing outcomes of 165 individuals using the LS/CM, LSI_R and HCR-20 found that sentencing outcomes were associated with risk assessment scores. ([Jung et al., 2015](#)).

Other Considerations

- When testing the internal consistency of the LS/CMI, [Gordon, Kelty and Julian \(2015\)](#) looked at how well items in the LS/CMI correlated with the overall score. Twelve items were removed as a result of this test. Further, it was highlighted that five of the items within the LS/CMI could be considered 'double-barrelled' questions: for instance, asking about both youth and adult criminal history in the same item. It is, thus, suggested that these items are separated to allow them to be adequately measured in this sample (Australian individuals aged 18-67 completing community-based sentences).
- The clinical/professional override feature was the focus of [Guay and Parent's \(2017\)](#) study of 3646 individuals who offended in Quebec. In 144 of 3646 cases, the clinical override was used to reduce the level of recidivism; whilst in 93 instances, the measure was used to increase it. It is concluded that the 'upward overrides' (i.e. increased level of threat) had greater predictive accuracy than 'downwards overrides' to a lower risk level. On this basis, it is recommended that further research is carried out on protective factors and the situations allowing for a 'downward override.' The results also showed that the clinical override feature decreased the predictive accuracy of the LS/CMI, apart from in cases of convictions for new crimes.
- [Wormith, Hogg and Guzzo \(2012\)](#) found that when assessors applied the override this tended to be to increase rather than decrease risk level. It was found that this reduced the predictive validity of the scale by excessively increasing risk. For instance, 263 individuals who committed sexual offences were initially in the low risk category and were thereafter overridden to medium, high or very high risk. Despite this, they actually recidivated (24.2% for medium risk; 19% for high risk; 5.9% for very high risk) at a lower rate than individuals deemed to be at low-risk (31.7%).
- Following on from this and similar findings from other studies, [Wormith and Bonta \(2018\)](#) highlight cautious use of the professional override. It is recommended that any 'excessive' use – defined as more than 5% of cases – should be quickly addressed to avoid prediction of recidivism being compromised.

- The LS/CMI does not measure religiosity (religious belief or feeling) or spirituality. A study by [Bhutta and Wormith \(2016\)](#) in the highly devout country of Pakistan added a measure of religiosity to the LS/CMI to test if the inclusion of this improved the predictive accuracy of the instrument. They, however, concluded that the addition of this to the instrument is unlikely to improve its ability to predict recidivism.
- The LS/CMI was pilot tested as the LSI-OR (Andrews, Bonta and Wormith, 1995) for a number of years prior to its publication in 2004.
- Recent independent research (not author-affiliated) has been conducted on LS/CMI. These include a study investigating the predictive validity of gang and non-gang members ([Guay, 2012](#)); an examination of the need principle and effect of treatment on change of LS/CMI scores ([Holliday, et al. 2012](#)); and a comparison on LS/CMI of psychopathic and neuropathic (characterised by frontal lobe deficits and psychosis) individuals who committed homicide ([Gilligan and Lennings, 2011](#)).

Name of Tool	Level of Service Inventory Revised (LSI-R)
Category	General Risk Assessment (Validated)
Author / Publisher	Andrews and Bonta
Year	(1995)

Description

- LSI-R is a 54-item actuarial tool of the individual's attributes and their circumstances. It is designed to assess criminogenic risk and identify the needs of those who have offended ([Watkins, 2011](#)).
- Information is collected via a semi-structured interview, a review of case records and collateral verification ([Wilson et al., 2016](#)).
- The tool centres on the principles of risk, need and responsivity, maintaining that those who are at high risk of reoffending should receive higher intensity interventions, supervision and monitoring ([Watkins, 2011](#)).
- Thirty-four items are subdivided across ten subsections. The total score is used to calculate recidivism risk, categorised as either 'minimum,' 'medium' or 'maximum.' Subscale scores are used to identify criminogenic needs ([Watkins, 2011](#)).
- In addition to recidivism, composite scores help to predict parole outcomes and the presence or risk of institutional misconduct ([Wilson et al., 2016](#)).
- Normed on North American prison, parole and probation populations.

Age Appropriateness

16+

Assessor Qualifications

Assessors must possess advanced training, certification and experience in psychological assessment or a related discipline, or satisfactorily complete a training course certified by the publishers. Can be used by a large range of professionals including social work and probation services.

Strengths

- Ability to discriminate risk across various outcome measures such as spousal abuse recidivism ([Hendricks et al., 2006](#)).
- Provides structured professional decision-making in a way that is comprehensive and consistent regardless of the case presented ([Campbell et al., 2009](#)).
- Both criminal history and the needs are captured with the tool. There is also an override feature to allow for the exercising of professional judgment to be exercised ([Wilson et al., 2016](#)).

- [Wilson and Stevenson \(2017\)](#) claimed that the semi-structured interview component of the instrument is a helpful framework for treatment and supervision, since it addresses learning, behavioural and developmental issues.
- The LS instruments are based on ‘General Personality and Cognitive Social Learning theory,’ which is a *general* theory of criminal conduct entrenched in social learning perspectives ([Wormith and Bonta, 2018](#)).

Empirical Grounding



- The LSI-R is supported by and reflective of three primary sources of information: (1) prior literature on recidivism, (2) professional opinions of probation officers and (3) social learning theory of criminal behaviours (Andrews and Bonta, 1995: 1).
- The subscales reflect the main risk factors identified in the research literature ([Andrews and Bonta, 2010](#)).
- Subject to a number of meta-analyses ([Olver et al., 2014](#))

Inter-Rater Reliability



a) UK Research

- [Palmer and Hollin \(2007\)](#) - inter-rater agreement levels of 95% for females.
- [Hollin and colleagues \(2003\)](#) found a 90% agreement rate in a sample of males.

b) International Research

- [Dahle \(2006\)](#) found excellent inter-rater reliability generating an ICC value of .93 in a sample of German individuals who had offended.
- [Lowenkamp et al. \(2004\)](#) - moderate to high levels of agreement observed across all ten subsections ranging from 61.5% to 97.7%.
- [Andrews \(1982\)](#) - excellent inter-rater reliability coefficients ranging between .80 to .99.
- [Persson et al. \(2017\)](#) found that the inter-rater reliability for the LSI-R was excellent (ICC=.92).

Validation History

General Predictive Accuracy



a) UK Research

- [Raynor and Miles \(2007\)](#) - predictive accuracy ranging from 65.4% to 71.6%.

	<ul style="list-style-type: none"> • Raynor (2007) - LSI-R presented ability to discriminate between reconvicted individuals who received a fine and those serving community/probationary sentences. • Hollin and Palmer (2006) found a moderate correlation between the LSI-R composite score and reconviction status.
b) International Research	<ul style="list-style-type: none"> • Duwe and Rocque (2016) administered the LSI-R to 26,000 prisoners in Minnesota for the time period of 2003 to 2011. The results gave an AUC of 0.628, providing moderate support for the LSI-R's ability to assess need. • In a study of 828 prisoners in Midwest of the United States, the LSI-R was able to predict recidivism (Smith et al., 2014). • A study in Australia found that the LSI-R yielded an acceptable level of reliability, with internal consistency estimates in the range of 0.59 to 0.784 (Watkins, 2011). • Campbell French and Gendreau (2009) - the LSI-R displayed one of the largest mean effect sizes in predicting violent recidivism ($Z+ = .28$). • A study by Lowenkamp et al. (2009) found moderate correlations between both re-arrest ($r = .36$) and re-incarceration rates ($r = .33$) and the LSI-R composite score. • Manchak et al. (2008) - the LSI-R yielded an AUC value of .73 for both general and violent recidivism. • Dahle (2006) - the LSI-R achieved moderate accuracy in violence prediction over a 10-year period (AUC = .65) in a sample of Germans.

Validation History	
Applicability: Females	
a) UK Research	<ul style="list-style-type: none"> • Raynor and Miles (2007) - for females in England and Wales ($n = 163$) the LSI-R mean score = 21.2, % correctly predicted = 65%. • Palmer and Hollin (2007) found that for female prisoners in England and Wales ($n = 150$) the LSI-R mean score = 23.0. There were significant differences between male and female scores on seven subscales, but not in

	<p>the overall score. Scores significantly predicted reconviction and time to reconviction. The composite score correctly classified 74%, with 79.7% correct classification for those not reconvicted and 64.9% for those who were convicted.</p>
b) International Research	<ul style="list-style-type: none"> • An Australian study found that the correlations between criminal history items and recidivism rates decreased in magnitude and significance when the LSI-R was applied to females. The author posited that the LSI-R subscales may not be suitable for fully assessing the criminogenic needs of females who offend (Watkins, 2011). • Hogg (2011) found the LSI to be gender neutral. • Manchak et al. (2008) - the LSI-R attained excellent predictive accuracy in relation to recidivism in a sample of female who offended (AUC = .77). • In a meta-analysis by Smith and colleagues (2009), it was found that the LSI-R demonstrated a correlation of $r=.35$ for recidivism in females. • Vose et al. (2008) - the LSI-R was found to be a valid predictor of recidivism in females, achieving a composite score of 71.4% accuracy.

Validation History	
Applicability: Ethnic Minorities	
a) UK Research	None at present.
b) International Research	<ul style="list-style-type: none"> • Hsu, Caputi and Byrne (2010) - the LSI-R demonstrated small correlations with recidivism in a sample of male and female Australian Indigenous individuals ($r_s = .12$ and $.16$ respectively). Indigenous individuals were found to score consistently higher on every item of the LSI-R. • Fass et al. (2008) - inconsistent validity with ethnic minority groups. LSI-R had better predictive accuracy with Caucasians (80.4%) and Hispanics (82.4%) than African Americans (43.4%). • Schlager and Simourd (2007) - few statistically significant correlations between LSI-R composite scores and recidivism amongst ethnic minority groups.

- [Ostermann and Salerno \(2016\)](#) applied the LSI-R to 9454 individuals in New Jersey to gauge its validity in predicting recidivism within a year of their release from prison. It was found that the LSI-R displayed low capacity for distinguishing between recidivists and non-recidivists when applied to Black males.

- A study by [Watkins \(2011\)](#) found that the discriminatory power on the LSI-R were very low for those with Aboriginal/Torres Strait Islander status in a sample of Australian individuals

- [Chenane et al. \(2015\)](#) examined the predictive validity of the LSI-R in 2778 male prisoners in the Midwest of the United States across White, Black and Hispanic ethnic groups. Results indicated that the LSI-R was better-suited to predicting institutional misconduct for White prisoners than the other two groups. It was suggested by the authors that the tool is modified to adhere to the risks and needs of Black and Hispanic prisoners.

- Applying the LSI-R to 95 clients within a mental health jail diversion program, [Lowder et al. \(2017\)](#) determined that the LSI-R showed weak predictive validity for African Americans than Caucasian clients. Moreover, the risk estimate was found to under-classify African Americans for the moderate risk category; whilst over-classifying them for high risk.

- Research by [Lowder and colleagues \(2019\)](#) suggested that there was no racial bias in the LSI-R. Analysis focused on 11792 probationers in Kansas (74.7% White and 25.3% Black). Risk classifications and total scores produced similar levels of predictive accuracy between the two groups.

- A meta-analysis of 32 articles and 12 data sets was undertaken to examine whether the LSI-R was applicable to Aboriginal individuals. Results indicated that all of the Central Eight risk/need factors were predictive of general and violent recidivism for Aboriginal individuals. Some of the factors demonstrated significantly better predictive validity for non-Aboriginal individuals: criminal history, alcohol/drug and antisocial pattern ([Gutierrez et al., 2013](#)).

Validation History

Applicability: Mental Disorders



a) UK Research	None at present.
b) International Research	<ul style="list-style-type: none"> • Harris, Rice and Quinsey (1993) found large weighted correlations ranging between .43 and .53 between items in the LSI and violent recidivism in a male psychiatric sample. Recidivists also tended to attain significantly higher scores on the tool than non-recidivists. • A study assessed 193 detainees who were undergoing a forensic psychiatric investigation in Stockholm. The predictive validity of the LSI-R was medium, generating an AUC of .70 (Persson et al., 2017).

Contribution to Risk Practice

- The LSI-R has the ability to create awareness of a number of static and dynamic risk factors pertinent to the individual's general risk of recidivism. Information obtained through the LSI-R can inform the level and focus of monitoring and supervision strategies.
- The tool can aid on-going evaluation of an individual's risk of reoffending and their criminogenic needs.

Other Considerations

- Fewer validation studies conducted with other populations such as ethnic minority groups and mentally disordered individuals.
- Requires refresher training - experience and training in the LSI-R can affect the reliability of the instrument ([Lowenkamp et al., 2009](#)).
- The tool is a quantitative survey of risk-need factors that are supported by research, professional opinion and social learning theory on criminal behaviour. It is not a comprehensive measure of mitigating and aggravating risk factors related to risk practices for offending (Andrews and Bonta, 1995).
- The LSI-R should be completed using information obtained from interviews with the individual and other collateral sources of information.
- The score of the LSI-R was found to correlate with the HCR-20^{v3} and the SAPROF at a considerable rate; although the correlations between the risk or protection categories were poorer ([Persson et al., 2017](#)).

Name of Tool	Level of Service Inventory-Revised: Screening Version (LSI-R:SV)
Category	General Risk Assessment (Validated)
Author / Publisher	Andrews and Bonta
Year	1995

Description

- The LSI-R:SV is an 8-item actuarial screening tool derived from the LSI-R. It encompasses seven key risk factors: criminal history, criminal attitudes, criminal associates, personal/emotional, employment, family and substance abuse.
- Similar categorisation of risk as observed in the LSI-R. High composite scores may warrant further analysis from the full LSI-R or LS/CMI assessment.
- Normed on Canadian institutionalised and probation populations.

Age Appropriateness

16+

Assessor Qualifications

Similar specifications as with its predecessor, the LSI-R.

Strengths

- Ideal for use when a complete LSI-R assessment may not be feasible, due to time constraints or insufficient staff resources. It is estimated to take between 10 and 15 minutes to administer.
- The LSI-R:SV can assist in prioritising cases for further intervention including assessment.

Empirical Grounding



- The LSI-R:SV is supported by and consistent with ‘...general personality, social psychological theory of criminal behaviour and the LSI-R items are consistent with an empirical body of literature and theory...’ (Andrews and Bonta, 1998:1).

Inter-Rater Reliability



a) UK Research

None at present.

b) International Research

- [Walters \(2011\)](#) - an estimate of inter-rater reliability from a random sample of 17 participants revealed an ICC of .71.

	<ul style="list-style-type: none"> Using a selection of 25 cases, Livingston et al. (2015) found that the LSI:R-SV had an ICC of 0.79.
--	---

Validation History	
General Predictive Accuracy	
a) UK Research	None at present.
b) International Research	<ul style="list-style-type: none"> Walters and Schlauch (2008) - the LSI-R:SV demonstrated moderate predictive accuracy in relation to recidivism in a male prison sample with AUCs for (1) official records of at least one officially reported incident (.63), (2) official records of at least one 'severe' incident (.62) and (3) self-reported incidents (.69). Yessine and Bonta (2006) compared 256 flagged individuals with 97 high-risk violent ones. High-risk violent individuals were found to scored lower on the LSI-SV. Examining the predictive accuracy of the 256 flagged individuals sample of the LSI-SV resulted in statistically significant results for all types of recidivism bar sexual recidivism. The AUCs generated were 0.68 for any recidivism, 0.67 and 0.63 for violent and non-violent recidivism respectively; sexual recidivism yielded an AUC of 0.53. The authors caution that the lack of predictive power in relation to sexual recidivism may be due to the relatively low base rate.

Validation History	
Applicability: Females	
a) UK Research	None at present
b) International Research	<ul style="list-style-type: none"> Lowenkamp et al. (2009) - the LSI-R:SV was not able to discriminate across the female risk categories of low, moderate and high. They recommend further larger sample research with subpopulations such as women.

Validation History	
Applicability: Ethnic Minorities	
No Empirical Evidence Available.	

Validation History

Applicability: Mental Disorders

a) UK Research

None at present.

b) International Research

- In a sample of 208 mentally ill individuals, the LSI-R:SV predicted recidivism with moderate accuracy (AUC) for the following: (1) any new offence (.67), (2) for non-violent new offences (.65) and (3) for violent new offences (.60) ([Ferguson et al., 2009](#)).

- [Thomas et al. \(2009\)](#) found the LSI-R:SV composite score generated moderate accuracy in predicting recidivism in a sample of forensic psychiatric patients (AUC = .72).

- The LSI-R:SV was applied to patients in a forensic psychiatric hospital in Australia to determine its scope to measure aggression risk. The results only showed a weak association between total scores and inpatient aggression, indicating that clinical factors pertaining to aggression should be incorporated into decision-making ([Daffern et al., 2005](#)).

- [Livingston and colleagues \(2015\)](#) conducted a retrospective review of health records for 250 probationers with mental disorders. Predictive accuracy using the LSI-R:SV was better for criminal justice contact and violent behaviour with AUCs of .61 and .67 respectively. The predictive power was less for non-compliance and psychiatric adverse event with AUCs of .58 and .55 respectively.

Contribution to Risk Practice

- The LSI-R:SV can aid the assessor in identifying some static and dynamic risk factors pertinent to the individual's likelihood of reoffending.
- The tool is useful for a brief scan of the main risk factors.
- The tool can alert assessors to the need to conduct a more thorough assessment.

Other Considerations

- Some research has found that the LSI-R:SV does not discriminate between those at moderate and high risk ([Lowenkamp et al., 2009](#)).
- The effectiveness of the LSI-R:SV for screening the offending population is based on preliminary and limited evidence ([Lowenkamp et al., 2009](#)).

- Assessors should note that this tool is a screening version of the full assessments (i.e. LSI-R, LS/CMI) and is not a comprehensive measure of risk and need factors.
- The LSI-R:SV should be completed using information obtained from interviews with the individual and other collateral sources of information.

Name of Tool	Offender Assessment System (OASys)
Category	General Risk Assessment (Validated)
Author / Publisher	Home Office
Year	2002

Description

- OASys is an actuarial risk and needs assessment tool used by the prison and probation services in England and Wales.
- The OASys is composed of 14 subsections and generates a summary risk score in order to assess likelihood of reoffending and risk of harm to self and others.
- In August 2009, the OASys General reoffending Predictor (OGP) and the OASys Violence Predictor (OVP) were introduced, and the old OASys score was discontinued. The August 2009 update also introduced 'layered OASys,' with Basic, Standard and Full assessments of similar structure but different length becoming available (Howard, personal communication, January 2013).
- The OGP and the OVP predict "the likelihood of nonviolent and violent proven reoffending respectively" by combining information on identified static and dynamic risk factors ([Howard, 2011: i](#)).
- An electronic version (eOASys) was introduced in 2005.

Age Appropriateness

18+

Assessor Qualifications

OASys assessments must be completed by prison or probation staff who possess the necessary knowledge of behaviours of those who offend. Continued refresher training on the administration and scoring of this tool is recommended.

Strengths

- OASys includes a section dedicated to assessing the suitability of interventions.
- OASys also incorporates a self-assessment component that allows the individual to record their views on their own risk/needs.

Empirical Grounding



- OASys is grounded in the 'what works' evidence base as per risk-need-responsivity principles with regards to reducing reoffending ([Moore and Howard 2015](#)).

•The manual states that the measure’s development was founded from prison and probation effective practice guidelines and from empirical grounding of the LSI-R and the Assessment Case management and Evaluation (ACE) (Home Office, 1999).

Inter-Rater Reliability	
a) UK Research	<ul style="list-style-type: none"> • Debidin (2009) - moderate inter-rater reliability was found in three case studies (ICCs ranged from .56 to .65). • Morton (2009) and Debidin (2009) - most reliable items in the OASys were: accommodation, lifestyle/associates, drug misuse, Education, Training and Employability, Relationships, Emotional Well-being and Attitudes. The least reliable items were: Financial Management, Alcohol, Thinking and Behaviour and Risk of Serious Harm.
b) International Research	None available at present.

Validation History	
General Predictive Accuracy	
a) UK Research	<ul style="list-style-type: none"> • Howard and Dixon (2012a) found that changes in OVP scores between the initial and final assessments, significantly predicted re-offending in a sample of Welsh individuals. • Debidin (2009) - the OASys achieved moderate to high AUC values for different types of offending ranging from ‘Homicide and assault’ (.66) to ‘Weapons Possession’ (.74). • Howard (2009) - the OASys achieved moderate predictive accuracy. Accuracy of the instrument improved when used with the OVP and the OGP. The AUC values improved to 80% for non-violent offending and 76% for violent re-offending compared to 76% and 68% obtained from the OASys scores alone. • Howard et al. (2006) - 26% of persons rated as ‘low likelihood of reconviction’ were reconvicted within 24 months, compared with 58% assessed as ‘medium-risk’ and 87% assessed as ‘high-risk’.
b) International Research	None available at present.

Validation History

Applicability: Females										
a) UK Research	<ul style="list-style-type: none"> • Debidin (2009) – the OASys attained moderate to high AUC values of .72 to .81 for violent and non-violent offences in a female offending sample (n=1,585) 									
b) International Research	None available at present.									

Validation History

Applicability: Ethnic Minorities										
a) UK Research	<ul style="list-style-type: none"> • Predictive validity for black and ethnic minority groups was found to be lower in a study by Howard (2015a). • Debidin (2009) - low to high AUC values obtained for individuals of other ethnic minorities ranging from .57 to .75 for violent and non-violent offences. • Howard and Dixon's (2012a) study of the OVP recorded 7% non-white participants in the 2002/2004 cohort and 8% non-white participant representation in the 2004/2005 data set. 									
b) International Research	None available at present.									

Validation History

Applicability: Mental Disorders										
No Empirical Evidence Available										
<p>Howard and Dixon (2012a) report: "The present psychiatric treatment item seems crude but reliable: unlike other items in section 10 (Emotional Well-being), only basic information and training are required to score it. Most OASys assessments record little or no direct information on personality disorder, psychopathic personality features or active psychotic symptoms."</p>										

Contribution to Risk Practice

<ul style="list-style-type: none"> • The OASys has the ability to create awareness of static and dynamic risk factors related to the individual's risk of recidivism. It can also prompt further assessment of identified risk factors. • The OASys assessments, including the OGP and OVP scores, are summarised to inform Pre-Sentence Reports within the National Probation Service for England and Wales. • Many of the factors identified by the OASys can act as targets for treatment/change. • Factors included in the OASys can inform offence analyses and risk formulations.

- The OASys can contribute to risk management plans for more complex cases which require intensive monitoring and more detailed offence analyses. The tool contains a 'Risk of Serious Harm' section which allows the assessor to identify factors related to this construct.
- Based on a validation study of almost fifteen thousand individuals who had committed sexual offences, a sexual offending component the 'Sexual Predictor' has been added to the OASys in order to predict contact sexual reoffending ([Howard and Barnett, 2015](#)).

Other Considerations

- [Fitzgibbon and Green \(2006\)](#) and [Fitzgibbon \(2008\)](#) - concerns relating to the accuracy of the OASys in predicting recidivism in sub-groups of such as those with mental disorders and ethnic minorities. Other concerns regarding its utility in aiding parole decisions.
- [Morton \(2009\)](#) and [Debidin \(2009\)](#) - limited inter-rater reliability of some of the subsections on the OASys.
- Few validation studies published in peer-reviewed journals. Majority of validation studies conducted by the Home Office.
- OGP Version 2 (OGP2) and OVP Version 2 (OVP2) have been peer-reviewed and will be published in a forthcoming Ministry of Justice publication (Howard, in preparation). The OGP2 and OVP2 includes an 'offence-free time' component, which enables estimates to be made for those who have spent time in the community without reoffending, given that the likelihood of reoffending is greatest immediately after sentence ([Howard, 2011](#)). The compendium also includes validation evidence for the predictive accuracy of the OGP2 and the OVP2 in different offending groups according to age, gender and ethnicity.
- The compendium will also include a study constructing and validating a predictor of sexual reoffending, provisionally named the OASys Sexual reoffending Predictor (OSP) (Howard, personal communication, January 2013).
- The implementation date of OGP2, OVP2 and OSP has not been confirmed.
- [Howard \(2015c\)](#) had a series of recommendations with regards to the positive factors of the OASys: assessors should recognise the importance of both positive (personal strengths) and risk factors, something which could be highlighted during training; monitoring the recording of positive factors to ensure current ones are being maintained and to be aware of the development of others.
- The National Offender Management Service reserves Crown Copyright on OASys. Please contact mark.nickson@noms.gsi.gov.uk for details on licensing.

Name of Tool	Offender Group Re-Conviction Scale Version 3 (OGRS3)
Category	General Risk Assessment (Validated)
Author / Publisher	Howard and colleagues
Year	2009

Description

- The OGRS3 is an actuarial assessment tool that is used in conjunction with the OASys risk assessment by the National Offender Management Service (NOMS) (Home Office, 2002) in order to inform and improve the static/dynamic predictor found in the OASys ([Howard et al., 2009](#)). It was originally owned by the Home Office and was later transferred to Her Majesty's Prison and Probation Service in England and Wales ([Howard, 2018](#)).
- The OGRS3 contains items pertaining to the age at time of current caution, the type of offence, prior criminal history (including duration in years) and gender of the individual being assessed ([Stephens and Brown, 2001](#)).
- The tool is used in conjunction with the OASys: this is designed to assess how likely an individual is to reoffend, identify and classify offending-related needs ([Moore, 2015](#)). It can also be used in cases where the OASys has not been completed.
- The tool generates a probability of reconviction ([Stephens and Brown, 2001](#)).
- A fourth version of the tool has been developed and is currently awaiting release. This includes a predictor of non-sexual, violent recidivism (OGRS4/V) ([Howard, 2018](#)).

Age Appropriateness

18+

Assessor Qualifications

Qualified Probation Officer with the relevant training and experience.

Strengths

- Due to a reduction in the number of items from nine to six, OGRS3 can be scored more quickly and accurately than previous versions ([Howard, 2018](#)).
- It can provide a prediction of risk within a 1 to 2 year time period.
- It provides a gendered estimate of risk, calculating it differently for females and males ([Howard et al., 2009](#); [Howard, 2018](#)).

Empirical Grounding



- The OGRS3 is grounded in extensive Home Office Policy research (Kershaw, 1999, Independent Conference Paper) dating from the previous two versions of the tool ([Copas and Marshall, 1998](#); [Taylor, 1999](#)).
- The criminal history ‘copas rate’ is the most complex part of the OGRS based on two factors: the length in years of an individual’s known criminal career and their total number of convictions. The ‘copas rate’ of an individual is higher when they have more criminal appearances within a short ‘criminal career’ (i.e. from their first through to their current offending) ([Howard, 2018](#)).
- The developers explored previous research in the United Kingdom about gender, age, current offence and criminal history as significant predictors for recidivism. Subsequent versions were refined by testing the validity across different groups of individuals who have offended; this then led to the age/gender interaction in OGRS3 ([Howard, 2018](#)).

Inter-Rater Reliability

--	--	--	--	--	--	--	--	--	--

No Empirical Evidence Available.

Validation History

General Predictive Accuracy



a) UK Research

- [Wood et al. \(2015\)](#) noted a link between reoffending rates and the OGRS score. Thirteen percent of those with a ‘very low’ likelihood of reoffending went on to do so in comparison with the 67% of those ranked as ‘very high’ who recidivated.
- [Howard and Dixon \(2012b\)](#) - the OGRS3 attained moderate accuracy in predicting violent reoffending (AUC = .70) in a dataset of 49,346 assessments.
- [Wakeling et al. \(2011a\)](#) - significant differences in mean OGRS3 scores between the recidivist and non-recidivist groups (20.3 versus 9.6 respectively). OGRS3 obtained moderate to high AUC values with different groups ranging from .65 (those convicted of sexual offences) to .86 (those convicted of violent offences).
- [Howard et al. \(2009\)](#) - the OGRS3 substantially improved the prediction of ‘proven’ re-offending for all individuals (AUC = .80), compared with its predecessor, the OGRS 2 (AUC= .78). For prisoners only, the OGRS3 generated an AUC of .84 compared to the Sentence Planning Predictor (AUC <.83) (n = 71, 914).
- [Howard \(2018\)](#) found that the AUC was strong at .80 for when coding was carried out using centrally-held records

	and when the coding was completed by probation workers.
b) International Research	None available at present.

Validation History	
Applicability: Females	<input checked="" type="checkbox"/> <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> <input type="checkbox"/>
a) UK Research	<ul style="list-style-type: none"> • Debidin (2009) – the OGRS3 obtained moderate to high AUC values of .81 and .70 for non-violent and violent offending respectively in a female offending sample. • Howard (2009) reported OGRS3 provided more accurate predictions for females than the previous version.
b) International Research	None available at present.

Validation History	
Applicability: Ethnic Minorities	<input checked="" type="checkbox"/> <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> <input type="checkbox"/>
a) UK Research	<ul style="list-style-type: none"> • Debidin (2009) – the OGRS3 obtained moderate to high AUC values ranging from .64 for those convicted of violent offences who were of ‘Other’ ethnic origin to .75 for non-violent mixed-race individuals.
b) International Research	None available at present.

Validation History	
Applicability: Mental Disorders	<input type="checkbox"/>
No Empirical Evidence Available.	

Contribution to Risk Practice
<ul style="list-style-type: none"> • The OGRS3 has an ability to guide awareness of some static risk factors and can prompt further need for assessment of the risk of reoffending. • OGRS3 scores are used within the National Offender Management Service (NOMS) as part of the risk/needs/responsivity-based criteria for targeting of offending behaviour programmes. • Howard et al. (2009) maintained that the predictive performance of OGRS3 can be optimised by using the OASys or the Asset (Youth Justice Board, 2003) in conjunction with it.

- The Youth Justice Board is in the process of rolling out the use of version 4 of the OGRS for young people ([Moore and Howard, 2015](#)).
- Formal training for the OGRS is provided for prison staff members as part of the OASys assessor course ([Howard, 2018](#)).

Other Considerations

- Authors claim that OGRS3 can be used within the youth justice system; although there is no empirical evidence to date to support this claim.
- Few validation studies by independent researchers.
- Assessors should note that the OGRS3 is designed to be used in conjunction with the OASys; hence the observed limitations in its capacity to contribute to risk practices on its own.
- OGRS prediction scales are used as a base measure in a number of settings. For example, the Ministry of Justice is linking OGRS with the 'payment by results' scheme. In Wales OGRS was used to evaluate a mentoring scheme for ex-prisoners. ([Maguire et al., 2010](#)).
- OGRS3 does not have a component to capture violent recidivism. The fourth version of the tool addressed this gap ([Howard, 2018](#)).
- [Howard \(2015a\)](#) found that the OGRS3 accurately measured rarer types of serious offences. It was recommended that arson, kidnapping, blackmail, dangerous driving and racially aggravated offences should be added to the OGRS3.
- A note of caution is potentially an individual's score can fall when they receive a new conviction. This is in scenarios when the length of the 'criminal career' is longer than the number of convictions or when age increases and an individual goes up in an age band. The decrease in OGRS score is to reflect the effect of growing older and longer breaks between offences. This should be considered, however, in the context of the behaviours and circumstances of individuals ([Howard, 2018](#)).
- OGRS Version 4 (OGRS4) will be introduced in a forthcoming Ministry of Justice publication (Howard, in preparation). The publication date has not yet been confirmed. Preliminary research has found the OGRS4, consisting of general and violent reoffending models, significantly outperforms the third version ([Howard, 2015b](#)).
- The OGRS4 includes a separate predictor of violent recidivism as well as a predictor of general recidivism. Both remain based on static risk factors, though OGRS4 also includes violent offending history. The OGRS4 publication includes tests of validity by age and gender, including young people. OGRS4 includes an 'offence-free time' component, which enables estimates to be made for those who have spent time in the community without reoffending, given that the likelihood of reoffending is greatest immediately after sentence ([Howard, 2011](#)).
- The National Offender Management Service reserves Crown Copyright on OGRS versions 3 and 4. Please contact mark.nickson@noms.gsi.gov.uk for details on licensing.