

RMA

PILOT OF AN ASSESSMENT FRAMEWORK FOR INDECENT IMAGES OF CHILDREN (IIOC)

PUBLISHED 2025



**REDUCING
SERIOUS HARM**

CONTENTS

1. EXECUTIVE SUMMARY	3
2. PURPOSE OF REPORT	4
3. PROGRESS OF THE PILOT	4
4. THE ASSESSMENT FRAMEWORK.....	6
5. EVIDENCE SUPPORTING THE USE OF CPORT & CASIC.....	7
6. EVIDENCE AGAINST THE USE OF CPORT & CASIC	10
6.1 Research conducted during the Scottish pilot.....	11
7. EVIDENCE OF OTHER RISK TOOLS.....	15
8. SUMMARY OF EVIDENCE AND DEFENSIBLE USE OF TOOLS.....	18
8.1 Risk and issues.....	19
8.2 Wider practice issues.....	22
8.3 Options considered.....	24
8.4 Recommendations.....	26
9. REFERENCES.....	28
10. APPENDIX.....	33

1. EXECUTIVE SUMMARY

This report relates to the use of risk tools in Scotland with individuals convicted of indecent images of children (IIOC) offending and particularly the implementation of a pilot of an assessment framework that included the use of the Child Pornography Offender Risk Tool (CPORT) & Correlates of Admission of Sexual Interest in Children (CASIC). This pilot took place between June 2022 and September 2024. This report summarises the extent of the evidence relating to the CPORT, CASIC, and other related risk tools, as gathered from the literature, research, and from this pilot.

A number of factors detrimentally impacted on the pilot. The low number of assessments during the pilot was the most significant issue. After 15 months of the pilot (June 2022 to September 2023) there were between 29 to 34% of the projected number of assessments based on pre-pilot estimated figures, and these figures remained consistent for the duration of the pilot. This did not appear to align with data on recorded crimes, rates of IIOC offending, and a rise in individuals in prison for IIOC. Court backlogs, time taken to sentencing, use of diversion, and several process issues, were highlighted as possible influencing factors. The low number of assessments, alongside issues accessing current recidivism data, were significant barriers to the planned research on testing the predictive validity of the assessment framework (including the use of the CPORT & CASIC).

Based on a review of the literature and research evidence relating to the reliability and validity of risk tools for men convicted of IIOC offending, the **CPORT, RM2000/S, STABLE-2007, and ACUTE-2007 (when used with the STABLE-2007) are all considered to be defensible tools to use for assessing risk of any sexual recidivism or IIOC recidivism. The Levels of Service tools may be used to predict general recidivism.**

Several significant practice issues were elicited during the pilot. These were:

- Assessors recommending supervision as a conclusion to an assessment in order to reduce the time someone spends on Sex Offender Notification Requirements (SONR).
- Inconsistencies in the application of nationally agreed risk tools such as the SA07.
- Assessments having pre-determined outcomes based on the presumed outcome from a decision-maker.

A draft of this report was produced in June 2024 to inform national strategic discussions at Multi-Agency Public Protection Arrangements National Strategic Group (MAPPA NSG) in July 2024. The purpose of which was to decide and agree upon the outcomes of the pilot and future actions. Future options following the completion of the pilot were presented and discussed, each with their own considerations and based on the evidence included in this report. These broadly consisted of:

1. Alternatives to the planned research on predictive validity.
2. Extend the pilot to try and capture more data for research purposes.
3. End the pilot and incorporate learning into existing risk assessment processes.
4. Roll the framework out nationally across Scotland based on evidence beyond the pilot.

Based on developments in the evidence based regarding defensible risk tools related to IIOC offending and learning from the implementation of the pilot and the research conducted during it, it was recommended at MAPPA NSG that the pilot was ended and learning incorporated into existing risk assessment processes. This was agreed by MAPPA NSG.

2. PURPOSE OF REPORT

This report sets out the existing evidence regarding risk assessment of individuals involved in IIOC offences and details the learning gathered from the pilot via project oversight and research.

In order to achieve this purpose, the report covers the following:

- A summary of the project.
- A summary of the current evidence relating to the Child Pornography Offender Risk Tool (CPORT)¹ and the Correlates of Admission of Sexual Interest in Children (CASIC)².
- A summary of the research undertaken as part of the pilot.
- A summary of the current evidence relating to the use of other risk tools with individuals convicted of IIOC.
- Issues that arose during the pilot and considered relevant to the implementation of the framework with exploration of contributing factors.
- Wider practice issues emerging during the pilot with impact beyond the remit of the pilot.
- Outcomes from the pilot and implications for practice.

3. PROGRESS OF THE PILOT



The pilot of a new assessment framework for those convicted of offences relating to IIOC commenced in June 2022. The pilot was part of a long-term project the Risk Management Authority (RMA) undertook in collaboration with Scottish Government (SG) and community justice partner agencies, in relation to internet offending risk assessment and management. The project can be described as having four phases, detailed below. This report is focused on the implementation of the pilot between June 2022 and September 2024.

Phase 1: In 2015, the Joint Thematic Review of MAPPA in Scotland³ produced numerous recommendations. One of them was for additional guidance to be developed to enable social workers to better assess the risk posed by internet offenders. In response to this, the RMA undertook a literature review on internet offending which was published in November 2018.⁴ It was agreed with SG in September 2018 that following the publication of the literature review, the RMA would develop an assessment framework for assessing the risk of individuals convicted of internet offending, specifically those convicted of offences relating to IIOC and with a particular focus on trying to defensibly identify individuals who were more likely to move from non-contact (such as internet) to contact sexual offending.

1 Seto & Eke (2015)

2 Eke & Seto (2016)

3 <https://www.careinspectorate.com/images/documents/2830/Joint%20Thematic%20Review%20of%20MAPPA%20in%20Scotland%20-%202015.pdf> (HMICS & Care Inspectorate, 2015)

4 <https://www.rma.scot/wp-content/uploads/2018/08/A-Review-of-the-Risk-Posed-by-Internet-Offenders.pdf>

Phase 2: This involved developing an assessment framework informed by the findings from the literature review. Following discussions with SG it was initially agreed that this process could link in with the national revision of Moving Forward Making Changes (MFMC).⁵ Originally a development working group for the assessment framework was set up as a sub-group of the MFMC risk assessment workstream. However due to COVID-19 the MFMC workstreams were halted. To progress development of the framework, the RMA formed its own development working group, featuring multi-agency representation including SG, Justice Social Work, Police and Academia. The working group met over a 12-month period and developed an assessment framework to be applied by Justice Social Work at the court report stage (i.e. post-conviction and pre-sentence). See Appendix B for the process outline map of when the framework was applied.

Phase 3: Following development of the framework, it was agreed that it would be piloted across Scotland. A two year pilot was agreed with SG for the framework to be applied across local authorities in Scotland. The RMA developed a set of criteria that areas needed to evidence to become a pilot site (see Appendix C for the full set of pilot site criteria). Criteria included training and supporting staff, sharing data, supporting research projects and involvement in a working group. The RMA opened applications to all local authorities and areas were selected as pilot sites based on their ability to meet the criteria. Eight local authorities applied and were selected to be involved in the pilot: Glasgow, Edinburgh, South Ayrshire, East Lothian, South Lanarkshire and the Tayside Partnership which includes Angus, Perth & Kinross and Dundee. Following selection of pilot sites, training was provided to all those who would be involved. The training was delivered in two parts. Firstly there was training on the CPORT and CASIC by one of the tools developers, and secondly, the RMA delivered training on the application of the framework. The pilot then commenced in June 2022. Throughout the pilot a working group met regularly consisting of representatives from the RMA, SG, pilot sites (such as Justice Social Work), and Police Scotland.



As part of the pilot it was agreed that research would be conducted to explore consistency, practitioners' views on usability and validity. The outcomes of this research are also contained in the evidence section of this report, as well as the impact of a number of barriers on the intended research. To summarise, the research on inter-rater reliability (i.e. consistency of scoring across assessors) and perceived usability (i.e. focus groups with practitioners completing assessments) were both completed and published alongside this report. The findings of each are woven into this report. However, when the pilot was planned there had also been the proposal to conduct a predictive validity study (i.e. do the outcomes of the assessment identify individuals more likely to move from IIOC to contact offending). This was impacted on by a number of barriers which are detailed in the issues section of this report and it meant this study was not conducted.

Phase 4: The fourth stage of this project is the post-pilot phase. This will be informed by the outcomes, evidence and options detailed in this report. Progress so far is reported but has included additional guidance to those assessing individuals with IIOC offences at the court report stage, awareness sessions to practitioners on assessment in this area, and the publication of this report alongside the inter-rater and usability studies.

⁵ This has since concluded with the implementation of the Moving Forward 2 Change (MF2C) programme.

4. THE ASSESSMENT FRAMEWORK

The framework incorporated the use of the CPORT and CASIC risk tools.

The CPORT (Seto & Eke, 2015) is a risk assessment tool designed to predict any sexual recidivism (contact and non-contact sexual offences, as well as IIOC⁶ alone) among adult male individuals with a conviction for a IIOC offence. It is made up of the following 7 items:

- i. Offender age at time of the index investigation
- ii. Any prior criminal history
- iii. Any prior or index failure on conditions such as probation, parole or conditional release
- iv. Any prior or index contact sexual offence history
- v. Indication of paedophilic interests
- vi. More boy than girl content (≥51%) in the child pornography content
- vii. More boy than girl content (≥51%) in child nudity and other child content, excluding child pornography content

The CASIC (Eke & Seto, 2016) was developed to help score Item 5 (Indication of paedophilic interests) on the CPORT. This is because this item was found to be susceptible to faking or refusal to respond. The CASIC is made up of the following 6 items (with a score of 3 or more leading to Item 5 on the CPORT being scored as 'present'):

- i. Never married
- ii. Child abuse video(s)
- iii. Content included sex stories involving children
- iv. Evidence of interest in child pornography spanning a time frame of 2 years or longer
- v. Volunteered with children before or during the index offence
- vi. Engaged in sexual communication with a minor, or undercover police officer posing as a minor, online

The pilot assessment framework incorporated these risk tools and provided a structure to analyse the pattern, nature, seriousness, and likelihood of an individual's offending behaviour. This was very similar to the Initial Assessment (which incorporates the screening version of the Level of Service assessment tool plus an initial analysis of offending) used by Justice Social Work to inform assessments at the court report stage. However, a key difference related to the underpinning risk tools being the CPORT and CASIC. Assessors concluded the assessment by using all the information within their assessment and analysis to make a judgement on an individual's level of both antisociality and sexual interest in children.⁷ This informed a decision making process relating to what was the most appropriate and proportionate risk management response (routine, intensive, or heightened measures). As such the assessment did not conclude solely based on scores or levels from the CPORT and CASIC but rather on the basis of a wider analysis of offending.

6 Note that the CPORT authors use the term Child Pornography. Other writers use terms like Child Sexual Exploitation Material (CSEM), Child Abuse Material (CAM), etc. This report will use the term 'Indecent Images of Children' (IIOC) throughout.

7 Research has highlighted antisociality and sexual interest in children as relevant factors when considering whether someone may be of a higher likelihood to move from internet to contact offending (see RMA, 2018).

5. EVIDENCE SUPPORTING THE USE OF CPORT AND CASIC

The following section provides a summary of the existing literature and evidence that supports the use of the CPORT. It is provided in order of publication.

In 2015, Seto and Eke developed the CPORT in a Canadian study involving 266 adult males with IIOC offences and a 5 year follow up period. They identified 7 variables associated with a greater likelihood of any sexual recidivism and these became the CPORT items. By reporting Area Under the Curve (AUC)⁸ they found the CPORT score (accounting for one missing item) was a significant predictor of sexual recidivism (AUC = .74). This included examining a version of the CPORT that excluded several items (relating to child content other than IIOC, and admission/diagnosis of sexual interest in children). This compact version was also found to be useful in predicting recidivism (AUC = .73). They stated the CPORT could be useful for ranking individuals but encouraged caution and the need for larger sample sizes to increase reliability and credibility.

A 2017 study by Seto and Eke looked specifically at the CASIC with a sample of 60 individuals with IIOC offending. Their findings indicated the CASIC score had a strong relationship with admission of sexual interest in children (Item 5 on the CPORT) (AUC = .81). They concluded CASIC scores could be substituted for admission of sexual interest in children in risk assessment.

Pilon (2016) used a modified CPORT and applied it to 279 individuals with IIOC offending and with a 3 year follow up period. The modified version of the CPORT included; age based on remand or sentence start date (rather than age at time of investigation), prior history based on convictions (not charges), indication of sexual interest in children based on luring or sexual interference involving youth, and omitting the two child content items. Findings indicated moderate predictive accuracy for general recidivism (AUC = .68). This included the CPORT outperforming a general offending tool (Levels of Service) that it was compared to.

Black (2017) applied a short version of the CPORT (CPORT-SV – the same as the shortened version used in the 2015 Seto and Eke study) on 547 individuals in New Zealand. The average follow up period was 7 years 7 months. They found the CPORT-SV was significantly associated with all 4 recidivism outcomes: any (AUC = .80), any sexual (AUC = .77), sexual contact (AUC = .74) and IIOC (AUC = .77). They also tested the CPORT-SV supplemented with 1 item from a New Zealand applied tool called the Automated Sexual Recidivism Scale (ASRS: Skelton et al., 2006) and this improved the predictive accuracy in relation to IIOC recidivism in particular (AUC = .82).

Eke and colleagues (2019) used a sample of 80 men convicted of IIOC. They found similar results to Seto and Eke's development sample from 2015 but with slightly lower predictive accuracy for sexual recidivism (AUC = .70; compared to .74 in the 2015 study). They also combined this with the 2015 sample and found that the CPORT predicted any sexual recidivism (AUC = .72) and IIOC recidivism (AUC = .74). This included being significantly predictive of these outcomes for individuals with IIOC offences and no known contact offending. This study also found high inter-rater agreement (i.e. consistency of scoring on the tool from multiple people within the study) for both the CPORT and CASIC total scores. Overall they concluded that the results from this validation sample were consistent with the 2015 development sample.

A study in 2019 by Hermann and colleagues looked at how consistently the CPORT and CASIC could be scored or applied. It involved 27 probation officers completing an exercise, after being trained on the CPORT and CASIC, that involved rating multiple case studies. The analysis included measures of inter-rater reliability and percentage correct. The CPORT total score and individual items demonstrated good consistency in scoring. Additionally, the total mean score for both practice cases was within 1 point of the correct score and 85.2% were scored correctly. For the CASIC, the reliability for the items and total score

8 AUC = area under the curve. It is a measure of predictive accuracy that can range from 0.00 to 1.00, with .50 signifying chance prediction, and 0 and 1 signifying perfect negative and positive prediction, respectively. AUCs of .56, .64, and .71 are considered small, moderate, and large effect sizes, respectively (Rice & Harris, 2005).

were difficult to interpret due to low variability in the data.⁹ As such they relied on percentage correct statistics and found good consistency in the scoring of each CASIC item and moderate consistency in the scoring of the CASIC total score with 72.2% scored correctly across both practice cases. Overall they reported the reliability for the CPORT as very good and moderate to good for the CASIC.

Savoie and colleagues (2021) applied the CPORT on a Scottish population of 144 men with IIOC index charges. The CPORT significantly predicted any (AUC = .79), any sexual (AUC = .79), and IIOC recidivism (AUC = .75) in a 5 year follow-up period, including when one item could not be coded. They also tested the CPORT without the sexual interest in children item, and the non-IIOC online content items (like the shortened version used by the authors in their 2015 study) and found good predictive validity in relation to any recidivism (AUC = .82), any sexual recidivism (AUC = .80) and IIOC recidivism (AUC = .77). The findings included good predictive accuracy for the sub-group of individuals with IIOC but no contact offending, but the numbers were too small to test individuals with IIOC and contact offending. This study also included inter-rater reliability and they established substantial to almost perfect agreement across all but one of the CPORT items. The CPORT total score had excellent agreement. CASIC item scores all had between substantial to almost perfect inter-rater reliability, and the CASIC total score also had excellent agreement.¹⁰ They concluded that the CPORT was a good predictor of risk of reoffending and was a valid risk assessment tool to use with Scottish individuals with IIOC offending.

Soldino and colleagues (2020) examined the use of the CPORT in a sample of 304 men arrested in Spain for IIOC offences using a 5 year fixed follow up period. They found slight predictive ability of the CPORT for IIOC recidivism outcomes when the CASIC was used to replace Item 5 of the CPORT (AUC = .56). They found CPORT total scores slightly helped predict new IIOC offending amongst individuals with solely IIOC offending (AUC = .57) as did the CASIC total scores (AUC = .70). They suggested that this may support the use of the CASIC as a stand alone risk assessment tool for IIOC recidivism although further validation research was recommended.

Cohen (2023) explored the use of the CPORT in the USA. They used natural language processing and machine learning to extract information relevant to coding the CPORT from files.¹¹ The sample was 5,768 men on community supervision in the USA. They found moderate predictive accuracy of the CPORT for sexual recidivism (AUC = .62). Their results included evidencing several of the CPORT items being linked to sexual recidivism: age at index investigation, sexual interest in children, previous criminal history, failure on conditional release, and contact sexual reoffending. This study also tested whether the predictive capability of the CPORT could be enhanced by modifying the tool. This involved 2 other versions of a tool. The first was using all the CPORT items and some other risk factors generated from the artificial intelligence (AI) used to extract data (e.g. prior non-contact offending, and violent offending) regardless of their association to sexual reoffending. This produced slight predictive accuracy (AUC = .61). The 2nd version was only the CPORT items and other risk factors significantly associated with sexual reoffending. This was only slightly better (AUC = .65). Regarding this study it should be noted that the use of AI meant the coding did not fully comply with the CPORT manual and so it should be considered a modified or approximated version of the CPORT. For example, admission of sexual interest in children was extracted from self-report, not interviews with police, and gender preferences were based on self-report rather than objective information about the IIOC collection.

The CPORT has also been applied in Germany (von Franqué et al., 2023). They used it on a sample of 132 non-justice involved men self-reporting having accessed IIOC and voluntarily seeking treatment for sexual interest in children. They scored the CPORT based primarily on self-report information, so like the

9 Interrater reliability statistics that correct for chance agreement can produce very low, zero, or negative values when there is low variability in the data (Krippendorff, 2011). A reliability statistic like Krippendorff's tries to account for raters agreeing by chance (for example, everyone picked the same response). This element can potentially be impacted on by various features of these kind of studies, such as: low number of potential response categories (for example, yes, no, unknown), low number of case studies, case studies having the same 'correct' response on an item. All these features together can produce low reliability alphas despite high consistency, or conversely not produce the reliability statistic at all.

10 The following guidelines were used with range of values for level of agreement: 'Almost Perfect' = .81 – 1.00; 'Substantial' = .61 - .80; 'Moderate' = .41 - .60; 'Fair' = .21 - .40; and 'Slight' = .00 - .20 (Landis & Koch, 1977).

11 Natural language processing involves equipping computers with the ability to understand text and spoken words in the way humans do. It enables computers to process human language in the form of text or voice data to understand its full meaning (including things like intent and sentiment) (<https://www.ibm.com/think/topics/natural-language-processing>)

Cohen (2023) study it would have deviated from the coding manual (for example, there was no investigation related to the index offence). They found the CPORT score using the CASIC significantly predicted self-reported future contact sexual offending against a child (AUC = .69) and self-reported IIOC use (AUC = .63) during an average of 28 months at risk.

Helmus and colleagues (2023) published a meta-analysis involving the use of the CPORT. Across 5 studies and more than 1,000 participants they found that the CPORT predicted any sexual recidivism (AUC = .75) and IIOC recidivism (AUC = .66) in men with any IIOC offending. They also reported that the CPORT was slightly better at predicting sexual recidivism for individuals with IIOC and contact offending (AUC = .72) compared to those with exclusively IIOC offending (AUC = .68). Furthermore they found the CPORT was slightly better than the Risk Matrix 2000 (RM2000) for predicting any sexual recidivism with men convicted of IIOC offences, although the RM2000 was slightly better at predicting sexual offending in individuals with contact and IIOC offending. They noted overall that the CPORT does well at predicting sexual recidivism when looking at a broad sample of individuals with IIOC convictions, but it was more difficult when research looked at narrow sub-groups (for example, individuals with no other criminal history) because it became harder to distinguish risk.

Helmus and colleagues (2025) published a summary of the evidence relating to IIOC risk tools. In this they add the Cohen (2023) results to the 2023 meta-analysis findings from Helmus and colleagues. Although they do advise some caution due to the large sample size in the Cohen study and the slightly anomalous results. Adding the Cohen study reduced the overall predictive accuracy of the CPORT relating to sexual recidivism (from AUC = .75 to .68). Adding the Cohen study also significantly increased the variability across studies included in the Helmus meta-analysis because the effect size from Cohen's study was significantly lower than the other five studies. Helmus and colleagues document that this evidenced the Cohen study to be an outlier and suggested this may be due to the unconventional data collection method used to approximate CPORT scores. However they also suggest the Cohen study demonstrates the possibilities of using machine learning and language analysis techniques in risk assessment which may mean more efficient processing of large quantities of data and large sample sizes.

Helmus and colleagues (2025) concluded that **the CPORT is defensible to use when assessing risk of any sexual recidivism or specifically IIOC recidivism among adult men convicted of IIOC offences.**

CPORT Version 3 is likely to be published in 2025/26 and contain an expansion of coding rules, more usable recidivism estimates, and risk levels.

6. EVIDENCE AGAINST THE USE OF CPORT AND CASIC

The following section provides a summary of the existing literature and evidence that could be considered not to support the use of the CPORT and CASIC.

Pilon's (2016) findings may have indicated moderate predictive accuracy for general recidivism but not for IIOC, sexual or violent recidivism (which was attributed to low rates of sexual reoffending within the sample at 2.9%, differences in coding sexual interest in children, and the missing items).

Soldino and colleagues (2020) (see on 'evidence for use of CPORT/CASIC' for aspects of this study that supported the use of the CPORT and CASIC) found observed recidivism rates were much lower than the expected recidivism rates presented by the CPORT tool developers. For example the CPORT under-predicted IIOC recidivism for CPORT total scores of 0 but predicted 90% more IIOC recidivists than was actually observed for CPORT total scores of 1, and over-predicted almost 3 times the number of actual recidivists for CPORT total scores of 2. This suggested caution over using the norms from the tool developers within risk assessment. The effect size reported by Soldino and colleagues (AUC = .57) for CPORT total scores predicting IIOC recidivism was also small compared to other studies. Furthermore, when restricting the analysis to cases with one or no missing items the predictive ability was no longer significant. Predictive accuracy of CASIC total scores did not reach significance for predicting IIOC recidivism. Soldino and colleagues suggested their findings may have occurred due to a number of factors: very low IIOC recidivism rates, the sample being mostly made up of individuals with IIOC-only offending (noting the CPORT authors indicated the inability of the CPORT to significantly predict sexual recidivism outcomes for this subgroup), differences in the population itself in terms of arrested vs. convicted individuals, and differences in data quality and completion compared to the CPORT development sample.

Brown (2022) reviewed the application of tools for IIOC offending and summarised some potential issues. They commented on studies being well conducted and reported but with relatively small samples that sometimes overlapped. Furthermore most studies used follow-up periods of less than 5 years, with the argument being for longer follow-up periods and larger samples to gain greater validity. A number of the studies also examined 'any sexual offending' rather than differentiating between IIOC and non-IIOC offending. Additionally many studies examined men who received custodial sentences which may not be representative of all IIOC individuals (for example, Soldino and colleagues used 'arrested' individuals and found lower CPORT scores and lower rates of prior offending). Brown also commented that the existing studies have barely included women (3 studies have), none have included ethnic minority samples, mental health conditions or disabilities, and only one study has addressed IIOC use in individuals with Autism Spectrum Disorder (ASD) (Allely et al., 2019). Focus in the studies on officially recorded offences or information may also mean there is an underestimation of actual contact sexual offending. They also highlight the practical issue of certain CPORT and CASIC items being potentially difficult to code based on available information (for example, other child nudity content). Brown's review concluded that describing the CPORT as a 'validated' tool was debatable due to the issues highlighted here. It should be noted that Brown's review was published in 2022 and, as such, is prior to the publication of the Cohen (2023) and von Franqué et al., (2023) studies, as well as the Helmus and colleagues' studies (2023, 2025).

Cohen's (2023) study reported that raters had difficulty in coding the CPORT items that measured boy to girl content in the IIOC material (CPORT Item 6) and other child nudity content (Item 7). They were also unable to code most of the CASIC items. The combination of high rates of missing data and the relatively low rate of sexual reoffending in the pilot sample of 195 individuals meant that the pilot aspect of this study elicited moderate to poor predictive ability for the CPORT (AUC = .54). As described in the 'evidence for' section this was improved slightly through a change in method to enable a much larger sample size of over 5000 individuals.

Scurich and Krauss (2023) published an article detailing significant criticisms regarding the use of the CPORT and CASIC. They referenced the development study (Seto & Eke, 2015) and commented on the CPORT's inability to predict recidivism for individuals with solely IIOC offending. They also stated that combining the development and validation samples across studies does not constitute independent validation of the CPORT. Furthermore, Pilon's (2016) sample may have overlapped with the author's samples. They concluded on 3 major criticisms of the existing research (again noting that this article was published prior to the 2023 Cohen study, Helmus and colleagues meta-analysis in 2023, and the 2025 review by Helmus and colleagues). Firstly, like the Brown review, they comment on the small sample sizes with none of them reaching the potential benchmark of 100 recidivists (as stated by Vergouwe et al., 2005). Small sample sizes can make a statistically significant effect more likely to not be a true effect (Button et al., 2013). Secondly they criticised the extent of missing data, particularly around CPORT Item 5 (sexual interest in children), Item 6 (more boy than girl content in the IIOC), and Item 7 (other child nudity content). Thirdly they comment on potentially outdated samples in that many of the studies use retrospective samples which spanned periods from the early 2000s to around 2015. The notion here is that technology is changing fast and it being debatable whether the samples used were representative of current IIOC offending.¹² Scurich and Krauss concluded that the current evidence did not support the use of the CPORT and CASIC to inform legal proceedings in the USA.

It should be noted, when considering the use of the tools in the current Scottish pilot, that they were incorporated into a wider assessment framework where there was not a reliance on scores or any sort of risk level. Rather the focus was to inform an analysis of offending (i.e. pattern, nature, seriousness, and likelihood) and judgements around antisociality and sexual interest in children, which then underpinned recommendations and decision-making (i.e. consistent with the literature around relevant risk factors in relation to individuals moving from IIOC to contact offending).

6.1 RESEARCH CONDUCTED DURING THE SCOTTISH PILOT

As indicated previously the pilot commenced in June 2022 across 8 local authorities. It included 3 pieces of planned research: inter-rater reliability, focus groups with practitioners on usability of the framework, and predictive validity relating assessment outcomes to recidivism data. This section reports on the findings of the inter-rater and focus group studies. The predictive validity was not conducted as planned, with the barriers referenced in the risks and issues section later in this report.

Inter-rater reliability study (RMA, 2025a)¹³

As part of the pilot 79 individuals were trained on the CPORT, CASIC and wider assessment framework. Upon conclusion of the training individuals were asked to complete an exercise where they rated a case study on the CPORT and CASIC (provided by the CPORT authors). 60 of these case studies were returned. A follow up exercise was then conducted a year later with 2 more case studies. This led to 17 individuals completing the further 2 case studies, and 14 raters having completed all 3 case studies. Analysis was conducted to produce statistics for inter-rater reliability and percentage correct.¹⁴ Due to the time gap between the 1st case study, and the 2nd and 3rd case studies (note these were also significantly shorter in length than the 1st), analysis also provided a breakdown of the 1st case study on its own as well as across all 3 case studies.

To analyse inter-rater reliability, a statistic called Krippendorff's alpha was produced (Hayes & Krippendorff, 2007). It is used to assess the reliability of coding between two or more raters and is suitable for categorical data (for example, where answers are things like 'yes', 'no', etc.). It is particularly

12 For example, Steel and colleagues (2022) found that current IIOC individuals are significantly more mobile in their viewing habits than they were a decade ago. For example more individuals are viewing IIOC in hotels, from vehicles, at work, etc. They also found that every individual in their sample who viewed IIOC at work also viewed it at home. Mobile phone use had increased indicating portable viewing was an increasing trend. Both points were related to the increasing potential need for residential search warrants based on work-based viewing and the seizure of mobile phones. Steel and colleagues also suggested that increases in stress, social isolation and downtime following COVID-19 may have increased IIOC offending.

13 Full publication available here: <https://www.rma.scot/publications/exploring-the-inter-rater-reliability-CPORT-CASIC/>

14 The original plan had been to use Krippendorff's alpha to replicate the Hermann and colleagues (2019) study. However it became apparent that this was problematic due to the low variability in the data. As such percentage correct was also used for items and total scores.

useful because it ignores missing data, can handle various samples sizes, categories and numbers of raters, and can be used on any measurement level (nominal, ordinal, interval, ratio). Values range from 0 to 1, where 0 is perfect disagreement and 1 is perfect agreement. Krippendorff (2004) suggests $\geq .800$ is ideal with $\geq .667$ being the lowest acceptable point (p.241).¹⁵ This approach was taken at various points during this study: to analyse individual items and the total score on the CPORT and CASIC across case studies 2 and 3 (administered at the same time), and to do the same across all 3 case studies. For some items the K alpha could not be produced and for others it could not be interpreted. These findings were checked with other professionals with expertise in this statistic (see footnote 15 for a link to the work of Marzi and colleagues, 2024). The conclusion reached was that the statistical analysis was likely being influenced by a number of factors such as the low number of case studies, a lack of variation in the 'correct' response to some items across case studies (for example, if all 3 case studies had the correct score of '1' to the same item), and the statistic conflating minor deviations in raters with a lack of reliability. This paper does not go into each individual example of this, however the full publication explores this in more detail. For this reason the percentage correct scores were useful to interpret alongside K alpha scores as together they provided a fuller picture of reliability in some instances.

The key findings from the inter-rater study were as follows:

- On the 1st case study the percentage correct (i.e. they got the same score as the actual case study example) was between 94-100% across all 7 CPORT items, demonstrating high consistency in correct scoring between raters. Similarly, the CASIC items also demonstrated high consistency in correct scoring between raters with all values above 90%. Across both tools there was some variance across individual items.
- For case studies 2 and 3, percentage correct scores indicated generally high levels of agreement across CPORT items. Over three quarters of individuals got the correct CPORT total score. Krippendorff's alpha indicated excellent reliability for CPORT total score and Item 6 (More boy than girl content in IIOC material). The K alpha for other items was influenced by the factors outlined above and explored in detail in the full publication. Dovetailed with percentage correct the only items where consistency of raters drops below a high threshold were Items 3 (Any prior or index conditional release failure) and Item 5 (Indication of paedophilic or hebephilic interests).
- Conversely, for case studies 2 and 3, the CASIC total score evidenced issues in terms of percentage correct, with only 50% of raters providing the correct rating (note this was better for Case Study 2). Inter-rater reliability analysis of CASIC total score via K alpha also indicated poor reliability. This was underpinned by significant issues with inconsistencies on Item 5c (IIOC material: stories/text) and 5d (Evidence of interest in IIOC spanning two or more years). This appeared, at least in part, to be influenced by inconsistency in scoring of certain items across one but usually not both case studies. For example, Item 5c was more inconsistently coded for Case Study 3. Whereas Item 5d was more inconsistently coded on Case Study 2.
- When analysing inter-rater reliability across all 3 case studies, the CPORT Item 7 (More boy than girl content in child nudity and other child material) maintained perfect agreement in terms of percentage correct. Items 1 (Offender age at time of the index: 35 or younger), 2 (Any prior criminal history), 4 (Any prior or index contact sexual offence history), and 6 (More boy than girl content in IIOC material) evidenced extremely high agreement. Items 3 (Any prior or index conditional release failure) and 5 (Indication of paedophilic or hebephilic interests) were a bit lower than the other items but still evidenced high levels of agreement. Overall the CPORT total score was correctly provided by 76.2% of raters across the 3 case studies. K alpha also indicated excellent inter-rater reliability for CPORT total score and Items 2, 4 and 6 (a slight improvement than when applied solely to case studies 2 and 3).

15 See Marzi and colleagues (2024) for further guidance on K alpha – <https://www.k-alpha.org/>

- However, across the 3 case studies, there was evidence of poor inter-rater reliability for the CASIC. At an item level this consisted of perfect agreement across Items 5a (Never married) and 5b (IIOC material: stories/text), fairly high agreement on 5e (Volunteering in a position with high access to children) and 5f (Engaging in online sexual communications with a minor or a police officer posing as a minor), and then poorer agreement on Item 5c (IIOC material: stories/text) and 5d (Evidence of interest in IIOC spanning two or more years). The K alpha indicated issues in reliability across Items 5c, 5d, 5e, and 5f. Overall CASIC score was correctly provided by 71.4% of raters across the 3 case studies. However the spread of responses, in particular for case studies 2 and 3 on Items 5c and 5d, evidenced significant inconsistencies across raters.
- In summary Case Study 1 (administered straight after the training) appeared to be scored more consistently than Case Studies 2 and 3 (administered as follow-up exercise). However across all 3 case studies there were issues in the reliability of the CASIC, which appears to be underpinned by particular inconsistencies in the scoring of 2 items relating to IIOC content involving sex stories about children and evidence of interest in IIOC spanning 2 years or longer.

There were several limitations and contributing factors relevant to these findings. Firstly this piece of research was conducted as part of the running of the pilot and so the method had to align with project timelines and procedures (for example, training, retaining the same raters, the workload of a rater, etc.). This had implications on the level of rigour possible within the constraints of managing the pilot. Anecdotal feedback from working group representatives was that administration of case studies 2 and 3 may have been slightly rushed due to timelines and perceived need. This could have impacted on the outcomes reported above. The sample of 14 raters and 3 case studies was also a relatively small cumulative sample and so individual raters disagreeing, and where disagreements themselves weren't consistent, could have impacted on findings. It is worth reiterating that there were better results attained for Case Study 1 (i.e. percentage correct) which was done immediately after the training and was also more detailed. This potentially supports the importance of the training on the CPORT and CASIC and the potential need for refresher training in the context of either a pilot or when fully implementing similar projects in the future, particularly where application of the tool may be infrequent.

Focus group study (RMA, 2025b)¹⁶

This study aimed to explore the views of professionals applying the assessment framework to gain an understanding of the benefits and challenges of using this method. Seven focus groups were conducted across various local authorities involved in the pilot which also allowed for the exploration of differences in practice between local authorities. The study used a framework analysis methodology and identified the following themes and subthemes:

Theme 1) Assessment framework is effective to apply

- *Subtheme A) Evidence-based decision-making through the use of the framework*
- *Subtheme B) The assessment framework promotes confidence with decision-making*
- *Subtheme C) The assessment framework (including CPORT/CASIC) is easy to use*

Theme 2) Challenges in applying the assessment framework

- *Subtheme A) Limited experience of applying the assessment framework*
- *Subtheme B) Difficulties assessing sexual interest in children*
- *Subtheme C) Lack of information available from external sources*
- *Subtheme D) Information sharing influenced by the quality of external relationships*

Theme 3) Sentencing outcomes

- *Subtheme A) Professionals considering the impact of a sentence*
- *Subtheme B) Communicating the outcome of the assessment framework*
- *Subtheme C) Recommendation of the assessment framework does not align with sentence imposed*

The analysis of the focus groups identified that the assessment framework was considered to be an effective method for assessing risk in this population, resulting in evidence-based decision making and promoting confidence. However, participants identified challenges which included assessing sexual interest in children, communicating the outcomes of the assessment framework as well as practice related issues. The findings of this study underpinned a number of implications for practice across: better information-sharing, clarification around the impact of notification requirement periods vs. supervision recommendations, refreshed understanding and application of Risk Need Responsivity (RNR) principles, and the benefits of focusing service quality assurance and/or audit mechanisms on the relationship between assessment recommendations and sentencing outcomes.



7. EVIDENCE OF OTHER RISK TOOLS

This section summarises the existing literature relating to the use of other risk tools for individuals with IIOC offending. It includes risk tools commonly used with other types of sexual offending (i.e. RM2000, Static-99R, Stable-2007, Acute-2007, and OSP), a further tool developed for use with individuals with IIOC offending specifically (i.e. ERICSO), and a Levels of Service general offending tool that has been explored (i.e. LSI-OR).

RM2000

Wakeling and colleagues (2011) conducted an initial exploration of the RM2000 scales (designed to predict sexual reoffending in individuals with sexual and/or violent offending). They used a sample of individuals with online sexual offences. The RM2000 (sexual scale) evidenced good predictive accuracy, however it was found to over-predict reoffending in that the observed reoffending of the sample was much lower than the tool predicted. A larger sample size and longer follow up periods were recommended.

The RM2000 was adapted in 2017 with the application of IIOC cases in mind. This included the item on stranger victim not being based on IIOC images, the item on non-contact offending being scored only if there was offline sexual offences as well in order to capture dual offending, and the item on male victim adapting its scoring rules based on whether an individual searched for the content.

The 2023 meta-analysis from Helmus and colleagues found that the RM2000/Sex predicted sexual recidivism (AUC = .66) and IIOC recidivism (AUC = .67) for individuals with a history of IIOC offending. However they also found that the CPORT (AUC = .73) outperformed the RM2000 in predicting any sexual recidivism. Regarding prior concerns about the RM2000 overestimating the risk in IIOC offending, the 2023 study found that estimated recidivism rates were actually significantly lower than the observed recidivism rates (the likely explanation given related to differing methods of measuring recidivism across the 2011 and 2023 studies).

The review article published in 2025 by Helmus and colleagues provided an overall summary of the use of the RM2000 with men with online offences. They commented on the improvements made to the scoring manual in 2017 being logical and based on existing evidence regarding IIOC offending. They concluded that the RM2000/Sex was defensible to use to assess the risk of any sexual recidivism or specifically IIOC recidivism among adult men convicted of IIOC offences. They referred to the evidence of moderate predictive accuracy, strong inter-rater reliability, and training and implementation resources being available. They concluded that although the tool is static and can't measure change over time, it can help organise treatment needs and make defensible decisions.

STATIC-99R

One study has examined the predictive accuracy of STATIC-99R with IIOC individuals. Eke (2022) examined it with 348 individuals from the CPORT development and validation samples, and 136 individuals from a second Canadian sample. In the sample of 348 individuals, the tool significantly predicted any sexual, contact sexual, and IIOC recidivism. However in the sub-group of 240 individuals with solely IIOC offending it did not significantly predict any sexual, contact sexual, or IIOC recidivism. For the sample of 136, the STATIC-99R significantly predicted any sexual recidivism, but again was not predictive for IIOC-only individuals.

Helmus and colleagues (2025) summarised the use of STATIC-99R with IIOC offending. Firstly the manual indicates the tool can be used with individuals with a IIOC offence only if they also have a charge or conviction for a Category A sex offence (contact and noncontact offending against identifiable victims). However, unlike the RM2000, the STATIC-99R manual does not offer guidance on coding cases where it is IIOC-only offending. Due to its coding rules, Helmus and colleagues commented that with IIOC exclusive offending it would effectively mean only 6 of the 10 items can be scored. They concluded that the STATIC-99R has partial evidence for defensible use with individuals with IIOC offending. It meets

minimum levels of predictive accuracy for any sexual, contact sexual, and IIOC recidivism, but with some significant caveats and limitations to the research. In particular the poor findings regarding IIOC-only individuals are concerning from a validity but also practical perspective, although the samples of these individuals used so far have been small and the research regarding this is somewhat preliminary at this time. Whereas with individuals with IIOC and other contact and/or non-contact offending the STATIC-99R appears to have some use. Beyond this and more generally in relation to assessing sexual offending the STATIC-99R has strong inter-rater and one of the strongest collections of evidence from research of any sexual recidivism risk tools.

Stable & Acute 2007 (SA07)

Brankley and colleagues (2021) conducted a meta-analysis that incorporated 72 individuals with IIOC offending as sub-group. They found the STABLE-2007 predicted risk of sexual recidivism in IIOC individuals (AUC = .94). They argued that their findings provided some initial evidence to support the use of the tool in this population. In particular they commented on the STABLE-2007 not having criminal history items, and as such the tool may be useful for individuals whose offending behaviour includes IIOC but not contact sexual offending.

Babchishin and colleagues (2023) conducted a meta-analysis which included 228 individuals with only IIOC offending, and 80 with IIOC and other offending. They found that the STABLE-2007 and ACUTE-2007 had similar predictive capabilities for IIOC groups (especially the IIOC-only group) compared to those with contact offending against children. For the individuals who had committed IIOC offences, both risk tools showed moderate to large associations with all the recidivism outcomes examined (any sexual, IIOC, any contact sexual, any violent, and any recidivism). In particular they found that emotional identification with children was one of the strongest predictors, and that significant social influences and capacity for relationship stability were particularly relevant to the IIOC groups. The authors suggested that the psychological and community adjustment factors targeted by the STABLE-2007 and ACUTE-2007 risk tools were relevant to the treatment and risk management of individuals who had only committed IIOC offences. They also discussed that these tools don't require information about the type of IIOC committed and they aren't limited to criminal history.

The review paper from Helmus and colleagues (2025) summarised the potential use of the STABLE-2007 and ACUTE-2007 tools. They referenced the growing evidence of their applicability with IIOC offending as well as the guidance now available to support coding these tools when using them with individuals with IIOC offending (for example, relating to victim access). They considered the STABLE-2007 to be defensible to use with individuals with IIOC offending. They still noted that the STABLE-2007 was designed to be combined with a static tool like the RM2000 (in order to produce recidivism estimates) and that currently no study had looked at this combination with IIOC offending. They also concluded that the ACUTE-2007 was defensible to use with IIOC offending in terms of assessing risk of any sexual recidivism or specifically IIOC recidivism. They reiterated that the ACUTE-2007 should be used in conjunction with the STABLE-2007. They did, however, state that no studies so far had looked at changes in risk over time with the IIOC group. This is important to keep in mind because this is one of the main benefits of the ACUTE-2007.

OSP/I

The OSP/I is a scale within a wider tool used in England and Wales called the OASys Sexual Reoffending Predictor (OSP). The OSP/I is based on indecent image offending history alone and is specifically designed to be used to predict IIOC reoffending. It replaced the use of the RM2000 in March 2021 in England and Wales. Helmus and colleagues (2025) explained that the OSP/I is less of a tool and more of a single item reflecting the individual's sexual offending history. It distinguishes between those with no IIOC offences, one IIOC offence, and multiple IIOC offences. In a sample of 2,728 men (593 with a IIOC conviction), the OSP outperformed the RM2000 with the OSP/I being the best predictor of indecent image reoffending. This version of the OSP/I had four risk levels (as opposed to three in another version) and significantly predicted (within an average 4.5 year follow-up) new IIOC convictions but did not significantly predict new contact sexual offences. There was no inter-rater reliability reported, however the single item approach, reliance on a review of sexual offence history, and scoring guidance availability,

means high reliability was expected. As indicated by Brown (2022) there is a need for replication with a non-prison sample.

Helmus and colleagues (2025) used the data from the OSP/I study to test a 3 risk level version and found a slightly significant effect for IIOC individuals (AUC = .59). They concluded that the OSP/I had some evidence and may be useful with men convicted of sexual offences to make a reasonable differentiation of risk of IIOC recidivism. However due to it being a single item it does not provide any utility in terms of informing approaches or decisions within assessment and risk management. Furthermore when applied to IIOC individuals it largely identifies the subgroup of individuals (<10%) who are higher risk for IIOC recidivism compared to the wider IIOC group. Its purpose is worth restating in terms of it being meant to be used to assess the risk of IIOC recidivism among all individuals convicted of sexual offences, not only among individuals convicted of IIOC offences. Due to the fact that an individual with IIOC offending can't be considered 'low risk' on the OSP/I it is argued that it is not in line with the risk principle of effective practice (Bonta & Andrews, 2023).

ERICSO

The Estimated Risk for Internet Child Sexual Offending (ERICSO) was developed by Garrington and colleagues (2022) and is a risk assessment tool for estimating the risk of IIOC reoffending and/or contact sexual offences. It was designed to be a structured assessment for individuals who have been convicted of at least one IIOC offence (including possession, downloading and trading IIOC). Individuals with contact offending must have at least one IIOC index offence. It is intended to support the identification of risk relevant treatment needs. The only study on the ERICSO focused on the development of the items included. Domains and items for the ERICSO were originally devised from a systematic review of literature relevant to IIOC reoffending. There was also a survey with 24 practitioners working with IIOC individuals and finally a user case study with an individual convicted of IIOC offences was also used to refine items. In their review of current tools, Helmus and colleagues (2025) commented that the ERICSO is a newer measure informed by literature and surveys and developed based on hypothesised factors relevant to IIOC convictions. They considered it to be more of a framework for data collection, and that it wasn't yet validated.

Levels of Service

The Level of Service Inventory (LSI) family of tools includes a number of different iterations which are all based around the 'Central Eight' risk factors for general recidivism (Bonta & Andrews, 2023). They are a widely used and established suite of tools for predicting general recidivism (Olver et al., 2014) and have evidenced consistently good inter-rater reliability (for example, Labrecque et al., 2018; Rocque & Plummer-Beale, 2004). In relation to IIOC, a version called the LSI-Ontario Revision (LSI-OR; Girard & Wormith, 2004) has been tested. It is very similar to the LS/CMI and includes 102 items across the Central Eight, as well as additional criminogenic needs, and responsivity factors (Wormith & Hogg, 2012). It can generate a risk/needs total score and a final risk level ranging from very low to very high risk. It includes the option for a practitioner to override the risk level, however this has been shown to consistently reduce predictive accuracy, particularly when used with individuals convicted of sexual offending (Wormith et al., 2012). Pilon's study included the LSI-OR in their examination of 279 men convicted of IIOC offences. They found a moderate correlation between the LSI-OR risk/needs total score and the CPORT. However the LSI-OR risk/needs score, initial risk level, and final risk levels did not significantly predict new IIOC offending, sexual offending, or violent offending. The risk/needs score (AUC = .63) and initial risk level (AUC = .62) did, however, significantly predict any recidivism. The final risk level after override did not. Helmus and colleagues (2025) concluded that LSI-related findings were preliminary at this time and more research was required. They stated the tools should not be used to predict sexual or IIOC recidivism with IIOC individuals. However they can be used to predict general recidivism, are useful for informing case management and treatment targets, but there should be caution around the use of overrides.

8. SUMMARY OF EVIDENCE AND DEFENSIBLE USE OF TOOLS

The pros and cons of each tool discussed above are presented in Table 1. There are no universally agreed required levels, criteria or thresholds of evidence to use a risk assessment tool. As such decisions on this must be based on multiple considerations and sources of guidance. Before offering a view on the defensibility of the tools included in this report, this section will begin with some of the relevant considerations regarding defensible use of a tool.¹⁷

Firstly decisions in a forensic context (for example, sentencing, progression, release, return to custody) will need to be made regardless of whether there is evidence available from a risk tool. Furthermore conducting and collating supporting research can take significant periods of time to establish and therefore decisions often need to be made using what is available. Research is also constantly indicating the need to change and update practices (particularly in the area of internet offending which is a rapidly evolving area). Therefore what constitutes acceptable evidence must be balanced against the alternative options (Helmus & Olver, 2023). An ideal risk assessment tool may be sought and desirable, however it may be considered reasonable to apply tools that are simply better than the alternative, particularly when that alternative is to apply unstructured clinical judgement which we know is prone to significant issues (for example, Grove & Meehl, 1996; Hart, 1998; Kemshall, 1996; Monahan, 1981). Selecting and using an appropriate risk tool is therefore about weighing up both the benefits and costs of both using or not using the tool.

To conclude on whether a tool is effective or not it is critical to remain mindful of its primary purpose; to assess risk of recidivism. As such the key criteria for effectiveness has to be the tools predictive accuracy (Hanson, 2022; Helmus & Babchishin, 2017). A valid benchmark is for a tool to be better than the aforementioned unstructured clinical judgement, which means looking for an AUC above .56 (or equivalent if using different statistical approaches), as this would be an improvement on unstructured clinical judgement. Consideration should also be given to the volume and quality of research, whether the tool helps inform assessment or treatment, whether it is based on the existing evidence of relevant risk factors, and whether it can be applied consistently.¹⁸ Factoring in these considerations and criteria enables for a more defensible decision regarding the appropriate use of tools. It is also the reason why often the use of multiple tools can be advantageous in terms of supporting different tasks (for example, to inform risk of recidivism or risk management approaches), or by adding incremental value to the overall risk assessment (Babchishin et al., 2012; Brankley et al., 2021). However understanding how to resolve different findings and discrepancies across multiple tools is still in early development (Thornton & Helmus, 2022).

Appendix A also provides a judgement on defensibility for each tool in relation to application with individuals with IIOC offending. This is based on the Daubert Criteria for Legal Admissibility. The Daubert standard is a set of criteria that can be used to determine the admissibility of expert witness testimony in federal court. It is used in the USA but provides a useful benchmark for the quality that evidence must be held to.¹⁹ This is explored in detail in Helmus and colleagues (2025) review paper. The Daubert criteria are:

- Whether the technique can be and has been tested
- Whether it has been subject to peer review and publication
- Its known or potential error rate
- The existence and maintenance of standards controlling its operation
- Widespread acceptance within the relevant scientific community

It is important to note that the research base in relation to IIOC risk assessment is still smaller compared to what exists for contact sexual offending. Larger sample sizes and longer follow-up periods are definitely desirable. Wider validation studies across individual variables (for example, gender,

17 Helmus and colleagues (2025) provide a more detailed and expansive explanation of the considerations that go into deciding whether or not to use a particular risk tool or not.

18 There are 2 tables provided at the back of the review paper from Helmus and colleagues (2025). These provide a really useful summary of the existing evidence regarding risk assessment with IIOC (<https://osf.io/preprints/psyarxiv/pngr9>). The criteria were developed by Helmus and Olver (2023).

19 For more information access <https://www.expertinstitute.com/resources/insights/the-daubert-standard-a-guide-to-motions-hearings-and-rulings/>

neurodiversity, etc.) are also desirable. However, as highlighted by Helmus and colleagues (2025), existing meta-analytic studies currently support the consistency of many relevant risk factors (Bonta & Andrews, 2023) including for sexual recidivism (Ahmed et al., 2023; Hanson & Morton-Bourgon, 2004, 2009; Helmus & Thornton, 2015; Helmus et al., 2022).

Based on the evidence, for men convicted of IIOC offending, the **CPORT, RM2000/S, STABLE-2007, and ACUTE-2007 (in conjunction with the STABLE-2007) are all defensible tools to use for assessing risk of any sexual recidivism or IIOC recidivism.** The Levels of Service tools may be used to predict general recidivism. It is important to keep in mind that the purpose of the assessment framework being piloted was specifically to support the assessment and identification of individuals who may move from internet to contact offending. Therefore, based on the literature, identification of antisociality and sexual interest in children were essential features of the pilot assessment framework. Furthermore the risk tools used were embedded within the Initial Assessment of Risk that incorporated an Analysis of Offending and could be used to inform supervision (if appropriate) with the individual (although a more detailed assessment will also be undertaken at that point). Considering all this there is logic in desiring both static and dynamic features to any assessment. The CPORT & CASIC fulfilled the static element, and the dynamic element took place following sentencing (depending on the outcome) should an SA07 then be applied (see wider practice issues section for related point on this and national practice). Taking all this into consideration and the evidence presented throughout the previous sections, it would be considered defensible for the CPORT & CASIC to be used in conjunction with the SA07. Another option could be the use of the LSI-R:SV (as is done in non-IIOC cases) alongside the SA07.

8.1 RISK AND ISSUES

This section details areas that arose during the course of the pilot which impacted on the delivery of key aspects such as research. Each issue is explained along with the potential underlying causes and the resulting impact on the pilot.

Low number of assessments

The pilot ran between June 2022 and September 2024. Data on the number of completed assessments was provided by pilot sites up until the end of September 2023 (see Table 1 below) to inform ongoing monitoring and evaluation. From a project management perspective the focus regarding data then switched to collection and analysis relating to the inter-rater reliability and focus group studies. Therefore the data below relates to a period of 15 months. Estimated numbers were also provided by pilot sites prior to commencing the pilot, noting that a potential underestimation was cited as more likely at the time due to the prior 2 years being impacted on by COVID. Based on the estimated numbers provided by pilot sites a 15 month projection was produced using both the lowest and highest estimations. As can be seen in Table 1 the estimated number of assessments at the 15 month point should have been between 317 and 367. The actual number of assessments was 107, between 29 to 34% – depending on lowest or highest estimations – of the original estimations.

Table 1. Estimated vs. actual number of assessments up until the end of September 2023 (rounded to the nearest whole case)

	Estimated number of Assessments over Two Years	15 month projection		Number of Assessments Completed until end of September 2023
		Lowest	Highest	
East Lothian	30 - 35	19	22	6
Glasgow	200 - 250	124	156	35
Edinburgh	40 - 60	25	38	13
South Lanarkshire	40 - 45	25	28	17
South Ayrshire	27	17	17	3
Tayside	170	106	106	33
Total	507 - 587	317	367	107

The low number of assessments was also inconsistent with wider evidence that indicated the prevalence of IIOC offending. The 'Recorded Crime in Scotland' report (the most recent one at conclusion of the data collection period in September 2023) detailed a 16% rise in IIOC offences between 2021-2022 and 2022-2023. Since 2013-2014 there has been an overall increase of 23%. Table 2 (see below) provides a breakdown of the different types of relevant offences across the past 10 years (Scottish Government, 2023a).²⁰ Whilst there have been fluctuations, the numbers for IIOC offending have risen since 2020. Justice Analytical Services (JAS) also publish monthly data which includes recorded sexual crime. The December 2023 edition referenced more online offending as a reason (amongst others) behind a long-term increase in recorded sexual crime (Scottish Government, 2023b).²¹ Scottish prison population statistics published in December 2023 also indicate a rise in the number of individuals given a custodial sentence for IIOC offences. There were increases in relation to 'indecent photos of children' from 63 in 2021-22 to 71 in 2022-23 (Scottish Government, 2023c).²²

Table 2. Recorded cases of sexual crimes

Sexual Crimes	2013-14	2014-15	2015-16	2016-17	2017-18	2018-19	2019-20	2020-21	2021-22	2022-23	% change 21-22 to 22-23	% change 13-14 to 22-23
Communicating indecently	605	906	1,180	1,166	1,222	1,373	1,470	1,543	1,608	1,518	-5.6	150.9
Communicating indecently	258	552	799	704	793	792	819	861	973	924	-5	285.1
Communicating indecently with older child (13-15)	244	189	182	237	203	290	317	313	268	309	15.3	26.6
Communicating indecently with young child (under 13)	103	165	199	225	226	291	334	369	367	285	-22.3	176.7
Indecent images of children	621	603	645	649	658	554	584	660	662	765	15.6	23.2
Communication Act 2003 (sexual)	-	-	-	270	301	199	223	354	315	354	12.4	-

Time to sentencing may also have been contributing towards the low number of assessments. Criminal justice monitoring data (Scottish Government, 2023d)²³ indicated that 46% of summary criminal cases were dealt with within 26 weeks from caution to verdict, and that the average Sheriff Court summary waiting period was 16 weeks between pleading diet and trial diet. This did not include information on solemn proceedings, which was reported in the JAS monthly data report (Scottish Government, 2023b).

Data in this report indicated a median time of 800 days from offence to verdict for sexual crimes. This may shed some light on the low number of assessments and may have had marked impact on the pilot. With this being a period longer than 2 years it may indicate why recorded crimes and police enforcement of IIOC appeared high but potentially this wasn't demonstrated and evidenced in the number of court reports completed by social workers for individuals with IIOC offending.

20 <https://www.gov.scot/publications/recorded-crime-scotland-2022-23/>

21 <https://www.gov.scot/binaries/content/documents/govscot/publications/statistics/2023/12/justice-analytical-services-jas-safer-communities-justice-statistics-monthly-data-report-december-2023-edition/documents/justice-analytical-services-jas-safer-communities-justice-statistics-monthly-data-report/justice-analytical-services-jas-safer-communities-justice-statistics-monthly-data-report/govscot%3Adocument/justice-analytical-services-jas-safer-communities-justice-statistics-monthly-data-report.pdf>

22 <https://www.gov.scot/binaries/content/documents/govscot/publications/statistics/2023/11/scottish-prison-population-statistics-2022-23/documents/report-22-23/report-22-23/govscot%3Adocument/Scottish%2BPrison%2BPopulation%2BStatistics%2B2022-23%2B-%2BAnalytical%2BReport%2B%25281%2529.pdf>

23 <https://www.gov.scot/publications/criminal-justice-monitoring-data-monthly-statistics/>

Related to the point above, court backlog was also raised as a possible contributing factor, in terms of how cases were processed or prioritised differently. Audit Scotland (2023) indicated that the backlogs of High Court cases and Sheriff solemn cases reached their highest levels in 2023. Projections suggested it may take until March 2025 for High Courts, and March 2026 for Sheriff solemn courts, to get back to normal operating capacity. Therefore, this backlog may have impacted on the number of assessments requested for IIOC-related offences. The working group also discussed whether backlogs may have been impacting on how cases were processed and prioritised, however this was difficult to ascertain. Police representatives on the working group reiterated that enforcement of IIOC offences were at an all-time high and as such the anticipation was that the number of assessments would, at some point, start to reflect that. It is therefore possible that the numbers of assessments will rise after the end of the pilot.

Increasing use of diversion from prosecution was also suggested as a potential contributing factor as this would have lessened the number of requests for assessments. The Justice Social Work Statistics in Scotland 2021-22 report (Scottish Government, 2023e)²⁴ detailed that the number of diversions from prosecution cases increased by 20% between 2020-2021 and 2021-2022. Data was not available for 2022-2023 at time of report drafting therefore it is unknown whether this has continued to increase. Within the Diversion from Prosecution: Joint Review publication (Scottish Government, 2023f),²⁵ it was noted that the Crown Office and Procurator Fiscal Service (COPFS) also holds data on diversion from prosecution however this was not published. Therefore, it is unknown whether there has been an increase since 2022 and if so, how much of an increase there has been.

A potential shift in offending trends may also be having an impact on the pilot. Pilot sites noted an increase in the request for court reports for online communications offences rather than IIOC which meant the pilot assessment framework wasn't suitable. Furthermore observations included seeing more instances of a mix of offences rather than solely IIOC. Whether this was due to a change in offending trends or an impact of justice processes such as the court, this was difficult to accurately ascertain and certainly to attribute causality to.

Other reasons provided were more practical and process-driven. Firstly, requests may have come through for reports where the framework could have been applied but due to not having the opportunity to use it before, social workers opted for familiarity in terms of following normal processes. Several working group representatives did support their staff to undertake refresher training as the training videos were recorded and remained available. However, it is anticipated that this will have only impacted a small number of cases and therefore does not account for the significant difference between estimated and actual number of assessments. Secondly, working group representatives noted there may have been instances where the framework could have been applied but for whatever reason, it was missed. This may have been related to lack of confidence, familiarity or awareness. However, the impact of this on reducing the numbers of assessments is again thought to have been minimal.

The consequences of the low number of assessments were significant. Planned research regarding predictive validity (i.e. Can the CPORT & CASIC predict recidivism? Do the judgements on antisociality and sexual interest in children within the framework relate to risk of recidivism?) required a large sample. The estimated numbers of between 507-587 would have been significant and could have allowed for a set follow up period (albeit that would have been limited). However the actual sample of 107 by September 2023 (following the same trajectories this would have become 171 at the 2 year point) was severely limiting and posed questions on the potential gain of investigating a small sample balanced against the necessary time and resources. This was also particularly important to consider in relation to establishing a follow-up period which meant that already small sample size would reduce further. For clarity, a small sample size increases the chances of false negatives or a 'missed detection' within the analysis. This is further compounded by IIOC being a low occurrence event, and so larger sample sizes help combat these problems and add power to any analysis.

The low number of assessments also meant there were social workers trained who had not had the chance to apply the assessment framework in their practice. Few social workers had the opportunity to use it multiple times. Table 3 below provides a breakdown of the numbers trained and how many had the

24 <https://www.gov.scot/publications/justice-social-work-statistics-scotland-2021-22/>

25 <https://www.gov.scot/publications/joint-review-diversion-prosecution/pages/1/>

chance to use it (as of the end of September 2023). It is important to keep in mind that there will have been some staff turnover in this period and newer members of staff coming in who may have done the training recently. However what also can't be ruled out is that this is another factor that influenced the inter-rater reliability study conducted as part of this pilot. In particular, the follow-up case studies may have included raters whose only exposure to the tool since the original training was this exercise (which may then have reduced consistency in scoring).

Table 3. Number of social workers trained across local authorities

Local Authority	Number Trained to Apply the Framework	Number Who Have Used the Framework in Practice
East Lothian	5	4
Glasgow	8	7
Edinburgh	21	8
South Lanarkshire	6	5
South Ayrshire	8	2
Tayside	25	Numbers vary per area
Total	73	26 confirmed

Recidivism data

In order to undertake predictive validity research, the RMA required recidivism data. Furthermore, as touched on above, follow-up periods are crucial in any predictive validity study. The reduced sample size (due to low number of assessments) impacted on this but another impact comes from recidivism statistics (published by JAS) often being published a year behind the current business year. For example, the data for 2023 may not be available until 2024. Overall a shorter, variable follow up period, and a reduced sample size all impacted on the quality of any research conducted and the level of meaning that could have been drawn from any findings.

Areas where the framework can't be used

There were important areas and individual characteristics where the assessment framework couldn't be used based on the criteria, and questions will remain around what the future of assessment in those areas looks like. The fast-changing nature of the internet also continues to pose challenges particularly in relation to offending and then crime identification, risk assessment and management. During the pilot there were fairly consistent questions regarding offences where images have been shared (through social media apps where images delete) but the individual had been convicted of an Online Communication offence. In these instances the framework could not be applied because the index offence had to have involved an IIOC offence. Furthermore, there remain questions regarding what to do in relation to assessing IIOC cases where it is a female, a young person, or across different responsivity factors (for example, autism).

8.2 WIDER PRACTICE ISSUES

During the course of the pilot a number of practice-related issues emerged with implications and locus beyond the pilot. As such they are detailed here for consideration at a national strategic level (e.g. such as at the MAPPA NSG) where multi-agency agreement can be reached on how to proceed and action change.

Registration periods for individuals convicted of sexual offending

This relates to periods of notification requirements impacting on and influencing the recommendations provided by social workers. This was initially raised and challenged during the training in 2022, however it continued to be raised throughout the course of the pilot and was identified as part of the focus

group study (RMA, 2025b). It appears that the length of time an individual will spend on notification requirements was a consideration when a social worker made their conclusions regarding disposal. This resulted in assessors recommending supervision to lessen the time spent subject to SONR. This had, and continues to have, implications in terms of principles of effective practice as it is contrary to the risk and need principles (Andrews & Bonta, 2010; Andrews et al., 1990). It is not proportionate to place people on supervision or interventions regardless of level of risk, and it can actually increase risk to over-manage people. For example, Bonta and colleagues (2000) found that low risk individuals receiving intensive levels of treatment evidenced higher recidivism rates (32%) than untreated low-risk individuals (15%). Furthermore, supervision should be purposeful and target relevant criminogenic needs, which is potentially not being done in these instances. It will also be significantly impacting on the resources of Justice Social Work. Discussions regarding this matter likely need to consider both the impact and length of notification requirement periods, as well as reminding practitioners of the principles of effective practice.

Another unintended consequence of this particular practice issue was that it likely confounded the outcomes in the pilot assessment framework. If social work were minded to recommend supervision in the majority of cases then this was contrary to the expected outcomes based on the literature around IIOC (i.e. the overwhelming majority being low risk and requiring minimal intervention). This would have almost certainly impacted on the validity research had it been conducted.

Appropriate and agreed application of risk tools

It was apparent from the pilot that the SA07 (a nationally agreed tool for those convicted of sexual offending) was applied at different stages of the court and sentencing process across different local authorities. This presented a particular challenge when considering the evidence in this report and trying to make recommendations regarding the appropriate use of tools for individuals with IIOC offending. Evidence suggests the SA07 may be defensible to use with this group, however should this be nationally agreed then this creates a discrepancy between the assessments done at the court report stage for IIOC and non-IIOC individuals. Based on level of concern and risk it is not advised that an IIOC individual could have (for example) a CPORT, CASIC and SA07 at court report, but a non-IIOC individual (i.e. a potentially 'riskier' person) has an LSI-R:SV and RM2K (with SA07 possibly used post-sentencing). The important point underpinning this is there is a need for national reconsideration regarding what tool to use, with who, and when, and for the outcome to be nationally consistent and proportionate.

Assessments pre-determining outcomes

Feedback during the pilot indicated that it was fairly common for the perceived decision or outcome of a Sheriff or Judge to influence the assessment and recommendations of a court report produced by a social worker (this was not specific to IIOC assessments). One example (amongst many) provided was there was no point in recommending a fine because the decision-maker would never go with this. This results in the assessor recommending what they think the sentencer wants. This undermines the purpose of risk assessment and introduces unnecessary bias. This may reflect a training need around the role and management of bias within risk assessment, as well as the roles and responsibilities of a risk assessor.

Over-reliance on tools

Observations during managing this pilot included a culture of over-reliance on risk tools and less discussion and emphasis on the importance of analysis. There was reluctance in some areas to stop using existing tools in addition to the use of the framework. At an operational level this created additional work on an already strained resource. Beyond that it is also likely to have included instances of inappropriate application of risk tools as well as disproportionate practice based on level of concern.

8.3 OPTIONS CONSIDERED

This section sets out the options that were considered by the MAPPA NSG in July 2024 when an interim version of this report was presented. These options drew from the evidence and findings across the various considerations set out in this report.

Inter-rater research

Due to the poor inter-rater reliability of the CASIC evidenced in the study conducted during the pilot, which is contrary to the more positive evidence from other studies, a further inter-rater exercise would have been justifiable. To improve on the rigour of the existing study this would have needed to involve administration of case studies within test conditions to ensure adequate time was taken and effort was applied. This could have included developing Scottish-specific case studies agreed with the CPORT authors (i.e. adapting them from existing Canadian examples).

This further inter-rater work would have been necessary if the decision had been taken to either extend the pilot or to implement the CPORT & CASIC as part of a national approach for assessing individuals with IIOC.

Predictive validity research

Due to the issues discussed above regarding the low number of assessments there would have needed to have been adjustments and/or alternative options considered for any predictive validity study (for example, does the assessment framework, including the CPORT and CASIC, accurately discriminate recidivism?). One option could have been to use a variable follow up period for the sample rather than a fixed one. In practical terms a fixed period would be, for example, taking everyone assessed in the first year of pilot and then a follow up period of 1 year (for timelines to be within the two year pilot). Whereas a variable period would involve taking everyone assessed in the first 18 months (for example) and linking to recidivism data at a certain point in time, meaning some people will have had a longer follow up period than others (i.e. those assessed nearer the beginning of the pilot).

The adjustments of follow up periods would still not have fully addressed the projection of sample size within the two year pilot. As a benchmark, to get to a sample size that matched the original development study (Seto & Eke, 2015) it would have required circa 250 assessments. At the rate of assessments stated in this report that would have taken 35 months from when the pilot started, and so this would have taken until May 2025. In addition to that, and even when using a variable follow up period, there would still needed to have been some additional time beyond that to allow the later assessments even just a minimum follow up period of six months. This would have taken the timeline to the end of 2025 based on projected assessment numbers. An estimated study completion time (to allow for analysis, and write up) would therefore have been around early to mid 2026. This also does not factor in the issues highlighted above around recidivism data potentially being a year behind, which could have pushed these projected timelines back even further.

An alternative approach to testing validity could have been to conduct it on a retrospective basis. The original proposal was to use a prospective sample in that individuals who were assessed as part of the pilot and then followed up on to consider recidivism. Whereas using a retrospective sample would have meant not using those assessed during the pilot but copying the approach taken in a number of other studies where the CPORT & CASIC were applied to a sample of historical individuals with IIOC convictions. This would have brought the advantages of producing essentially as large as sample as desirable (depending on how far you go back), being able to establish a fixed follow up period as desirable, and not necessarily being limited to pilot sites as it could have been done nationally. The disadvantages would have been that CPORT & CASIC would had to have been scored for each case. Furthermore the pilot did not plan for this piece of work and would have resulted in a significant dataset with increase burdens around data input and analysis. It is also worth reiterating that it would not be limited to the actual assessments conducted during the pilot.

If the pilot had been extended or rolled out nationally then some form of predictive validity study would have been explored with the recommendation of a long-term commitment to longer follow up periods (such as five years, 10 years, etc.). Related to this, the inclusion of the SA07 as an alternative or complementary tool for assessing IIOC individuals means that conducting a meaningful validation study could be explored as an alternative.

Pilot options

There were three options following the conclusion of the pilot in September 2024. These were: a) extend the pilot, b) end the pilot, or c) roll the framework out nationally. Each possibility included specific considerations.

a) Extend the pilot

This would have involved extending the pilot beyond the end date of September 2024. This would have allowed for an increase in assessment numbers on the assumption that factors like court backlog, time to sentencing, etc., were impacting the low numbers. This may have allowed for the predictive validity study to take place, although this was not guaranteed. It could have also permitted the completion of an additional inter-rater exercise.

b) End the pilot

This would have involved the pilot ending in September 2024 and the CPORT & CASIC not being implemented nationally. This would be based on emphasis being placed on the poor CASIC inter-rater reliability established in the pilot study, although the limitations of that study are worth keeping in mind as other studies have shown good reliability. Learning from the pilot, and the current literature on IIOC, could be integrated into existing practices (for example, updating court report guidance to incorporate the importance of factors like antisociality and sexual interest in children in order to inform sections like the Analysis of Offending). Based on existing evidence beyond the pilot, the use of the LSI-R:SV (i.e. current practice in non-pilot sites) could also be considered for use alongside the SA07, to try and assess the relevant domains of antisociality and sexual interest in children.

c) Roll the assessment framework out nationally

This would have involved expanding the use of the CPORT, CASIC and assessment framework to more or all local authorities in Scotland. This would have been based on the existing evidence on the use of the CPORT as presented in this report. It would be necessary that this option was complemented with a commitment to a further piece of inter-rater reliability work as well as long-term predictive validity (such as five years, 10 years, and so on). It could have incorporated formal review periods to allow for monitoring of emerging evidence and practices.

8.4 RECOMMENDATIONS

This section considers the evidence presented in this paper and the options detailed above to make recommendations regarding the RMA's view on the most appropriate course of action.

The RMA recommended that the pilot was concluded but the learning acquired during it (from the growing literature and evidence base, and the research conducted during the pilot) was incorporated into existing Scottish assessment approaches, practices, and guidance.

This recommendation was based on:

- The low number of assessments and the impact this had on the intended research and the resource commitment required in managing the pilot. Projections based on the rate of assessments detailed in this report would have taken the pilot into timelines way beyond those manageable from within the RMA's existing resource.
- Related to the above, the resource required to manage the pilot and the related components (such as working group, research, maintaining up-to-date knowledge, presenting at forums and conferences) was significant and less justifiable based on the low number of assessments.
- There was a growing research evidence base, since the pilot began, of the appropriateness and defensibility of using existing tools like the SA07.
- The LSI-R:SV could be used to assess general offending (i.e. to inform antisociality) and the SA07 as a sexual offending specific risk tool (i.e. supporting analysis of sexual interest of children). The CPORT and CASIC could also be used independently of the assessment framework.
- Placing some emphasis on the evidence of inter-rater concerns relating to the CASIC, although acknowledging the limitations of this study and the contrary wider evidence.
- Similarly, placing some emphasis on the there being at least one item on the CPORT that Scottish information collection systems do not support the rating of (i.e. Item 7 – more boy than girl content in other child nudity content).

This recommendation was presented to the MAPPA NSG in July 2024 where the decision was taken to end the pilot.

To support this outcome effectively the following additional recommendations were made:

- The RMA to lead on developing a chapter to be inserted into the existing Court Report guidance on assessing individuals whose offending is IIOC. This would align with the existing format of the guidance and be part of 'Specific Areas for Consideration'. It would incorporate key aspects of the assessment process and underpinning literature which align with the existing approach (e.g. incorporating antisociality and sexual interest in children into an Analysis of Offending). Note that since this recommendation was presented to MAPPA NSG (July 2024) this chapter has been drafted and is estimated to be published in 2025.
- There should be discussion and agreement of when the SA07 is applied nationally to ensure consistency in practice and proportionality across offence types (for example, contact to non-contact).
- More generally it may be beneficial for the appropriate national forums (for example, MAPPA NSG, SWS Standing Committee) to consider a review around what is the most reasonable level of assessment to be conducted at the court report to inform defensible decisions.

- To support effective use of the SA07 in relation to internet offending, consideration should be given to the SA07 training provided by Community Justice Scotland incorporating an internet offending case study. The RMA could provide support in the production of this case study.
- To support wider national understanding of assessing internet offending, the RMA would develop and run awareness sessions covering the key principles in this area. These sessions were delivered online in May 2025 to just under 500 staff members working in justice (including Justice Social Workers and Police staff).
- Should the decision be taken to focus practice around the inclusion of the SA07 then conducting a meaningful validation study would also relate back to the Scottish Government commitment to the Justice Sub-Committee 2 that the SA07 would be validated. This could involve collection and analysis of all SA07s conducted across police and social work (to provide an extremely significant sample size) with specific attention paid to individuals with internet offences. This could then be linked to recidivism data.
- The Practice Issues detailed in the corresponding section of the report (including consistent application of risk tools) may represent process and training deficiencies. They should be discussed at a relevant multi-agency national senior forum (such as MAPPA NSG, SWS Standing Committee) and taken forward as actions for change in order to improve national practices around risk assessment and application of principles of effective practice.



9. REFERENCES

- Ahmed, S., Lee, S. C., & Helmus, L. M. (2023). Predictive accuracy of Static-99R across different racial/ethnic groups: A meta-analysis. *Law and Human Behavior, 47*(1), 275- 291. <https://doi.org/10.1037/lhb0000517>
- Allely, C. S., Kennedy, S., & Warren, I. (2019). A legal analysis of Australian criminal cases involving defendants with autism spectrum disorder charged with online sexual offending. *International Journal of Law and Psychiatry, 66*, Article 101456. <https://doi.org/10.1016/j.ijlp.2019.101456>
- Andrews, D. A., & Bonta, J. (2010). Rehabilitating criminal justice policy and practice. *Psychology, Public Policy, and Law, 16*(1), 39.
- Andrews, D. A., Bonta, J., & Hoge, R. D. (1990). Classification for effective rehabilitation: Rediscovering psychology. *Criminal Justice and Behavior, 17*(1), 19-52.
- Audit Scotland. (2023). Criminal Courts Backlog. Scotland: Author. Available here: [Criminal courts backlog \(audit-scotland.gov.uk\)](https://audit-scotland.gov.uk)
- Babchishin, K. M., Hanson, R. K., & Helmus, L. (2012). Even highly correlated measures can add incrementally to predicting recidivism among sex offenders. *Assessment, 19*(4), 442- 461. <https://doi.org/10.1177/1073191112458312>
- Babchishin, K. M., Dibayula, S., McCulloch, C., Hanson, R. K., & Helmus, L. M. (2023). ACUTE-2007 and STABLE-2007 predict recidivism for men adjudicated for child sexual exploitation material offending. *Law and Human Behavior, 47*(5), 606-618. <https://doi.org/10.1037/lhb0000540>
- Black, C. (2018). *Predicting recidivism among an adult male child sexual abuse imagery offender population with the Child Pornography Offender Risk Tool Short Version (CPORT-SV): A New Zealand validation study* [Master's thesis, University of Canterbury]. UC Research Repository. <http://doi.org/10.26021/6729>
- Bonta, J., & Andrews, D. A. (2023). *The psychology of criminal conduct (7th ed.)*. Routledge.
- Bonta, J., Wallace-Capretta, S., & Rooney, J. (2000). A quasi-experimental evaluation of an intensive rehabilitation supervision program. *Criminal Justice and Behavior, 27*(3), 312-329. <https://doi.org/10.1177/0093854800027003003>
- Brankley, A. E., Babchishin, K. M., & Hanson, R. K. (2021). STABLE-2007 demonstrates predictive and incremental validity in assessing risk-relevant propensities for sexual offending: A meta-analysis. *Sexual Abuse, 33*(1), 34-62. <https://doi.org/10.1177/1079063219871572>
- Brown, S. J. (2022). Assessing the risk of users of child sexual exploitation material committing further offences: A scoping review. *Journal of Sexual Aggression*. Advance online publication. <https://doi.org/10.1080/13552600.2022.2104394>
- Button, K. S., Ioannidis, J. P. A., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S. J., & Munafò, M. R. (2013). Power failure: Why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience, 14*, 365-376. <https://doi.org/10.1038/nrn3475>
- Cohen, T. H. (2023). Building a risk tool for persons placed on federal post-conviction supervision for child sexual exploitation material offenses: Documenting the federal system's past, current, and future efforts. *Federal Probation, 87*(1), 19-31. <https://www.uscourts.gov/federal-probation-journal/2023/06/building-risk-tool-personsplaced-federal-post-conviction>

- Eke, A. W. (2022, October 26-29). *Predictive accuracy of the Static-99R in two samples of men adjudicated for child sexual exploitation material offending*. In K. M. Babchishin (Chair), *Risk assessment with unique groups: Child pornography and other non-contact sexual offenses* [Symposium]. Association for the Treatment of Sexual Abusers 41st Annual Research & Treatment Conference, Los Angeles, CA, United States.
- Eke, A. W., Helmus, L. M., & Seto M. C. (2019). A validation study of the child pornography offender risk tool (CPORT). *Sexual abuse*, 31(4), 456-476.
- Eke, A. W., & Seto, M. C. (2016). Scoring guide for the Child Pornography Offender Risk Tool (CPORT).
- Garrington, C., Kelty, S. F., Rickwood, D., & Boer, D. (2022). The development of the ERICSO – A proposed instrument for internet child abuse material offender assessment. *The Journal of Forensic Practice*, 24(4), 453-470. <https://doi.org/10.1108/JFP-07-2022-0033>
- Girard, L., & Wormith, J. S. (2004). The predictive validity of the Level of Service Inventory Ontario Revision on general and violent recidivism among various offender groups. *Criminal Justice and Behavior*, 31(2), 150-181. <https://doi.org/10.1177/0093854803261335>
- Grove W. & Meehl P. (1996) Comparative efficiency of informal (subjective, impressionistic) and formal (mechanical, algorithmic) prediction procedures: the clinical-statistical controversy. *Psychology, Public Policy and Law*, 2, 293–323. <https://doi.org/10.1037/1076-8971.2.2.293>
- Hanson, R. K. (2022). *Prediction statistics for psychological assessment*. American Psychological Association. <https://doi.org/10.1037/0000275-000>
- Hanson, R. K., & Morton-Bourgon, K. (2004). *Predictors of sexual recidivism: An updated meta-analysis* (User Report 2004-02). Public Safety and Emergency Preparedness Canada. <https://www.publicsafety.gc.ca/cnt/rsrscs/pblctns/2004-02-prdctrs-sxl-rcdvsmptd/2004-02-prdctrs-sxl-rcdvsm-pdtd-eng.pdf>
- Hanson, R. K., & Morton-Bourgon, K. E. (2009). The accuracy of recidivism risk assessments for sexual offenders: A meta-analysis of 118 prediction studies. *Psychological Assessment*, 21(1), 1-21. <https://doi.org/10.1037/a0014421>
- Hart S.D. (1998) The role of psychopathy in assessing risk for violence: conceptual and methodological issues. *Legal and Criminological Psychology* 3, 121–137. <https://doi.org/10.1111/j.2044-8333.1998.tb00354.x>
- Helmus, L. M., & Babchishin, K. M. (2017). Primer on risk assessment and the statistics used to evaluate its accuracy. *Criminal Justice and Behavior*, 44(1), 8-25. <https://doi.org/10.1177/0093854816678898>
- Helmus, L. M., Kelley, S. M., Frazier, A., Fernandez, Y. M., Lee, S. C., Rettenberger, M., & Boccaccini, M. T. (2022). Static-99R: Strengths, limitations, predictive accuracy meta-analysis, and legal admissibility review. *Psychology, Public Policy, and Law*, 28(3), 307- 331. <https://doi.org/10.1037/law0000351>
- Helmus, L. M., Eke, A. W., Farmus, L., & Seto, M. C. (2024). The CPORT and Risk Matrix 2000 for men convicted of child sexual exploitation material (CSEM) offenses: A predictive accuracy comparison and meta-analysis. *Criminal Justice and Behavior*, 51(1), 3-23. <https://doi.org/10.1177/00938548231208194>
- Helmus, L. M., Eke, A. W., & Seto, M. C. (2025). What risk assessment tools can be used with men convicted of child sexual exploitation material offenses? Recommendations from a review of current research. *Law and Human Behavior*. Advance online publication. [pre-print available online <https://osf.io/preprints/psyarxiv/pngr9>]

- Helmus, L. M., & Olver, M. E. (2023). *What evidence is needed to use a recidivism risk tool in court?* [Unpublished manuscript]. Department of Criminology, Simon Fraser University.
- Helmus, L. M., & Thornton, D. (2015). Stability and predictive and incremental accuracy of the individual items of Static-99r and Static-2002r in predicting sexual recidivism: A meta-analysis. *Criminal Justice and Behavior*, 42(9), 917-937. <https://doi.org/10.1177/0093854814568891>
- Hermann, C., Lucenti, L., & Eke, A. W. (2019). *CPORT and CASIC: Exploring consistency in scoring between probation and parole officers*. Ontario Ministry of the Solicitor General.
- His Majesty's Inspectorate of Constabulary in Scotland (HMICS) & The Care Inspectorate (2015). *Joint Thematic Review of MAPPA in Scotland*. Scotland: Author.
- Howard, P., & Wakeling, H. (2021). *Comparing two predictors of sexual recidivism: The Risk Matrix 2000 and the OASys Sexual Reoffending Predictor*. HM Prison & Probation Service, Ministry of Justice Analytical Series. <https://www.gov.uk/government/publications/comparing-2-predictors-of-sexualrecidivism-rm2000-and-osp>
- Kemshall H. (1996). *Reviewing Risk: A Review of Research on the Assessment and Management of Risk and Dangerousness: Implications for Policy and Practice in the Probation Service*. Home Office, London.
- Krippendorff, K. (2004). Reliability in content analysis: Some common misconceptions and recommendations. *Human communication research*, 30(3), 411-433.
- Labrecque, R. M., Campbell, C. M., Elliott, J., King M., Christmann, M., Page, K., McVay, J., & Roller, K. (2018). An examination of the inter-rater reliability and rater accuracy of the Level of Service/Case Management Inventory. *Corrections*, 3(2), 105-118. <https://doi.org/10.1080/23774657.2017.1323253>
- Marzi, G., Balzano, M., & Marchiori, D. (2024). K-Alpha Calculator —Krippendorff's Alpha Calculator: A user-friendly tool for computing Krippendorff's Alpha Inter-Rater Reliability Coefficient. *MethodsX*, 12, 102545. Available at <https://www.k-alpha.org/>
- Monahan J. (1981). *Predicting Violent Behaviour*. Sage Library of Social Research, Beverly Hills.
- Olver, M. E., Stockdale, K. C., & Wormith, J. S. (2014). Thirty years of research on the Level of Service scales: A meta-analytic examination of predictive accuracy and sources of variability. *Psychological Assessment*, 26(1), 156-176. <https://doi.org/10.1037/a0035080>
- Pilon, A. (2016). *The predictive validity of general and offense-specific risk assessment tools for child pornography offenders' reoffending* [Master's thesis, University of Saskatchewan]. University of Saskatchewan's Research Archive. <http://hdl.handle.net/10388/ETD-2016-01-2414>
- Risk Management Authority (RMA). (2018). *A review of the risk posed by internet offenders*. Risk Management Authority. Available at: <https://www.rma.scot/publications/a-review-of-the-risk-posed-by-internet-offenders/>
- Risk Management Authority (RMA). (2025a). *Exploring the Inter-Rater Reliability of the Child Pornography Offender Risk Tool (CPORT) and Correlates of Admission of Sexual Interest in Children (CASIC)*. Paisley: Author. Available at: <https://www.rma.scot/publications/exploring-the-inter-rater-reliability-CPORT-CASIC/>
- Risk Management Authority (RMA). (2025b). *Using framework analysis to explore views of practitioners applying an assessment framework piloted in Scotland with individuals convicted of Indecent Images of Children (IIOC)*. Paisley: Author. Available at: <https://www.rma.scot/publications/using-framework-analysis-explore-views-IIOC>

- Rocque, M., & Plummer-Beale, J., (2014). In the eye of the beholder? An examination of the inter-rater reliability of the LSI-R and YLS/CMI in a correctional agency. *Journal of Criminal Justice*, 42(6), 568-578. <https://doi.org/10.1016/j.jcrimjus.2014.09.011>
- Savoie, V., Quayle, E., Flynn, E., & O'Rourke, S. (2021). Predicting risk of reoffending in persons with child sexual exploitation material offense histories: The use of Child Pornography Offender Risk Tool in a Scottish population. *Sexual Abuse*, 34(5), 568-596. <https://doi.org/10.1177/10790632211047190>
- Scottish Government. (2023a). *Recorded Crime in Scotland, 2022- 2023*. Scotland: Author. Available here: [Recorded Crime in Scotland, 2022-23 - gov.scot \(www.gov.scot\)](https://www.gov.scot/Recorded-Crime-in-Scotland-2022-23)
- Scottish Government (2023b). *Justice Analytical Services (JAS): Safer Communities and Justice Statistics Monthly Data Report*. Scotland: Author. Available here: [Justice Analytical Services \(JAS\): Safer Communities and Justice Statistics Monthly Data Report \(www.gov.scot\)](https://www.gov.scot/Justice-Analytical-Services-JAS-Safer-Communities-and-Justice-Statistics-Monthly-Data-Report)
- Scottish Government (2023c). *Scottish Prison Population Statistics 2022-23: A National Statistics Publication for Scotland*. Scotland: Author. Available here: [Scottish Prison Population Statistics \(www.gov.scot\)](https://www.gov.scot/Scottish-Prison-Population-Statistics)
- Scottish Government (2023d). *Criminal justice monitoring data: monthly statistics*. Scotland: Author. Available here: [Criminal justice monitoring data: monthly statistics - gov.scot \(www.gov.scot\)](https://www.gov.scot/Criminal-justice-monitoring-data-monthly-statistics)
- Scottish Government. (2023e). *Justice Social Work Statistics in Scotland: 2021-22*. Scotland: Author. Available here: [Supporting documents - Justice Social Work Statistics in Scotland: 2021-22 - gov.scot \(www.gov.scot\)](https://www.gov.scot/Supporting-documents-Justice-Social-Work-Statistics-in-Scotland-2021-22)
- Scottish Government. (2023f). *Diversion from Prosecution: Joint Review*. Scotland: Author. Available here: [Footnotes - Diversion from prosecution: joint review - gov.scot \(www.gov.scot\)](https://www.gov.scot/Footnotes-Diversion-from-prosecution-joint-review)
- Scurich, N., & Krauss, D. A. (2023). Risk assessment of child-pornography-exclusive offenders. *Law and Human Behavior*, 47(4), 499- 509. <https://doi.org/10.1037/lhb000053>
- Seto, M. C., & Eke, A. W. (2015). Predicting recidivism among adult male child pornography offenders: Development of the Child Pornography Offender Risk Tool (CPORT). *Law and Human Behavior*, 39, 416-229. <https://doi.org/10.1037/lhb0000128>
- Seto, M. C., & Eke, A. W. (2017). Correlates of admitted sexual interest in children among individuals convicted of child pornography offenses. *Law and Human Behavior*, 41, 305- 313. <https://doi.org/10.1037/lhb0000240>
- Skelton, A., Riley, D., Wales, D., & Vess, J. (2006). Assessing risk for sexual offenders in New Zealand: Development and validation of a computer-scored risk measure. *Journal of Sexual Aggression*, 12(3), 277-286. <https://doi.org/10.1080/13552600601100326>
- Soldino, V., Carbonell-Vayá, E. J., & Seigfried-Spellar, K. C. (2020). Spanish validation of the Child Pornography Offender Risk Tool. *Sexual Abuse*, 33(5), 503-528. <https://doi.org/10.1177/1079063220928958>
- Steel, C. M. S., Newman, E., O'Rourke, S., & Quayle, E. (2022). Improving child sexual exploitation material investigations: Recommendations based on a review of recent research findings. *The Police Journal*. Advance online publication. <https://doi.org/10.1177/0032258X221142525>
- Thornton, D., & Helmus, L. M. (2022, June). Using multiple actuarial instruments to assess sexual recidivism risk. Online training for Society for the Advancement of Actuarial Risk Needs Assessment (SAARNA).

- Vergouwe, Y., Steyerberg, E. W., Eijkemans, M. J., & Habbema, J. D. F. (2005). Substantial effective sample sizes were required for external validation studies of predictive logistic regression models. *Journal of clinical epidemiology*, 58(5), 475-483. <https://doi.org/10.1016/j.jclinepi.2004.06.017>
- von Franqué, F., Bergner-Koether, R., Schmidt, S., Pellowski, J. S., Peterson, J. H., Hakak, G., & Briken, P. (2023). Individuals under voluntary treatment with sexual interest in minors: What risk do they pose? *Frontiers in Psychiatry*, 14, Article 1277225. <https://doi.org/10.3389/fpsy.2023.1277225>
- Wakeling, H. C., Howard, P., & Barnett, G. (2011). Comparing the validity of the RM2000 scales and OGRS3 for predicting recidivism by Internet sexual offenders. *Sexual Abuse*, 23(1), 146-168. <https://doi.org/10.1177%2F1079063210375974>
- Wormith, J. S., & Hogg, S. M. (2012). The predictive validity of Aboriginal offender recidivism with a general risk/needs assessment inventory. University of Saskatchewan. https://cfbsjs.usask.ca/documents/research/research_papers/LSIOR%20Aboriginal%20Paper%20w%20Abstract.pdf

10. APPENDIX

APPENDIX A: PROS AND CONS OF DIFFERENCE IIOC TOOLS

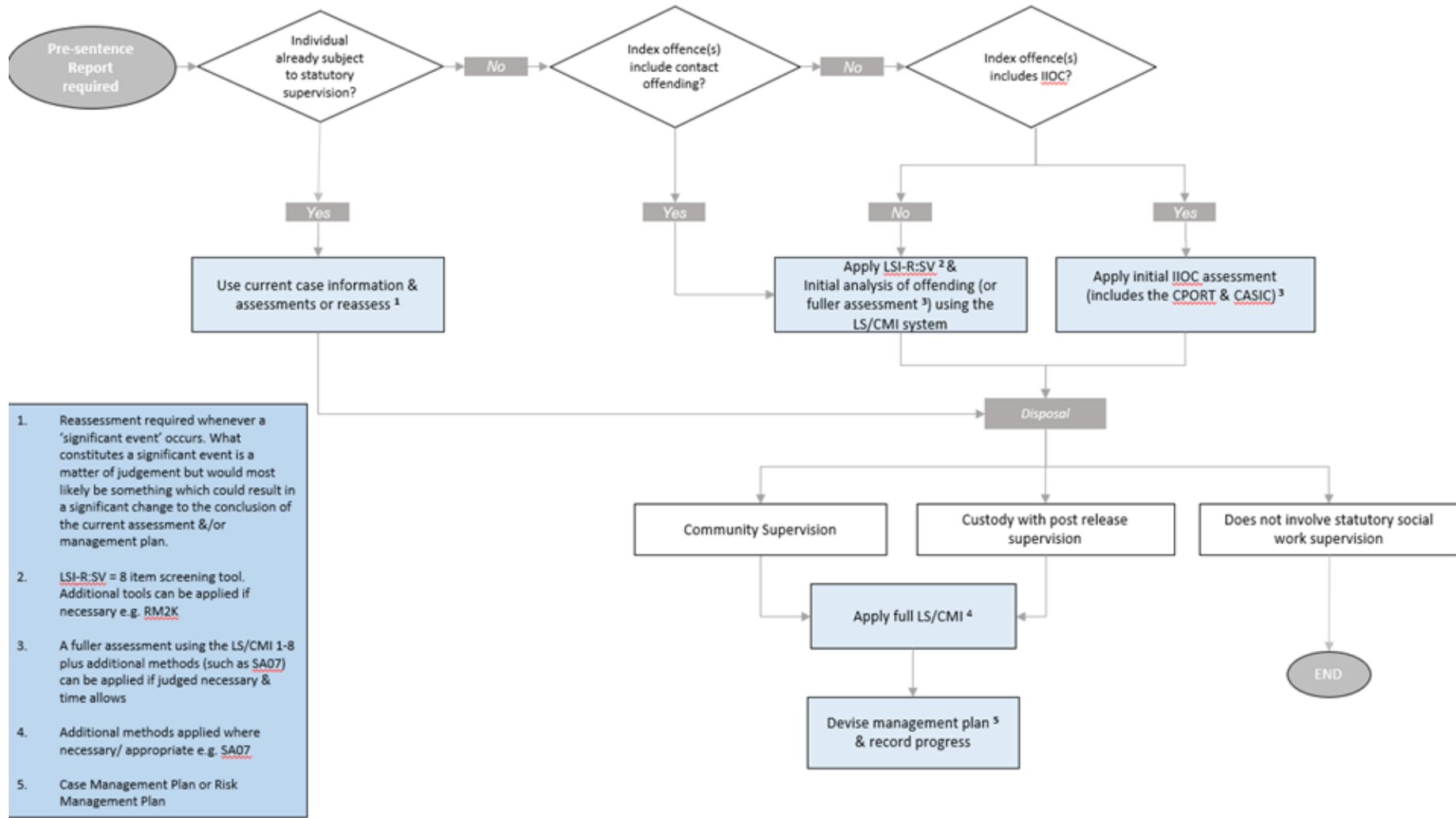
Tool	Pros	Cons	Defensible
CPORT & CASIC	<ul style="list-style-type: none"> • Items informed by literature and available information • Can account for some missing data • Moderate to large predictive accuracy (international and Scottish) • Largest pool of research of any IIOC tool • Predicts any sexual and IIOC recidivism • Short version also shown some predictive accuracy • Can be used to rank individuals • Items (particularly on CASIC) useful for risk management • International evidence of CPORT and CASIC inter-rater reliability • Scottish evidence of CPORT inter-rater reliability • Outperforms other tools in this space (e.g. RM2000) • Freely available and scoring guide supports coding • Helps assess antisociality and sexual interest in children 	<ul style="list-style-type: none"> • Items can be difficult to code depending on available information • Does not distinguish between types of recidivism • Issues in the pilot with CASIC inter-rater reliability • Need larger samples and longer follow-up periods • Does not assess changes in risk over time • Need for consideration of more individual variables such as gender and neurodiversity 	Yes
RM2000/S	<ul style="list-style-type: none"> • Moderate predictive accuracy • Strong inter-rater reliability • Predicts any sexual and IIOC recidivism • 2017 adaptations to items and guidance considered application to IIOC • Recent study found a relationship between RM2K/Sex and CPORT total scores • Can help inform treatment needs and decision-making 	<ul style="list-style-type: none"> • Some potential evidence of over-estimating IIOC reoffending • CPORT appears to outperform it • Static tool that can't measure changes in risk over time • Generic recidivism estimates that are not IIOC specific • Most of the support for RM2000 does relate to contact sexual offending 	Yes
STATIC-99R	<ul style="list-style-type: none"> • Predicted IIOC recidivism across individuals with sexual offending • Arguably meets minimum levels of predictive accuracy for any sexual, contact sexual, and IIOC recidivism • May be useful with dual offending • Strong inter-rater reliability • One of the strongest collections of evidence on sexual recidivism 	<ul style="list-style-type: none"> • Limited evidence of application to IIOC offending • Significant caveats and limitations to the research • The sole study found it did not predict recidivism in sub-group of IIOC individuals • No additional coding guidance for IIOC-only individual • Use with IIOC-only individuals would mean using only 6 out of the 10 items • Generic recidivism estimates that are not IIOC specific 	Partial

Tool	Pros	Cons	Defensible
SA07	<ul style="list-style-type: none"> Evidence is growing around moderate to good predictive accuracy with IIOC individuals 2023 meta-analysis indicates similar predictive capability for IIOC groups across STABLE and ACUTE Emotional identification with children, significant social influences and capacity for relationship stability may be particularly useful with IIOC Psychological and community adjustment factors may be particularly relevant to treatment and management of IIOC individuals Predicts any sexual and IIOC recidivism Don't require information about the type of IIOC committed and they are limited to criminal history Additional scoring guidance now available to support use with IIOC. Fairly robust in response to missing items The general use of the tool includes assessing changes in risk over time Currently implemented in Scotland with training and refresher training. 	<ul style="list-style-type: none"> Fairly small IIOC samples so far. Designed to be used with a static tool like RM2K which yet remains untested Studies haven't yet tested the use of the tool with IIOC individuals in terms of changes in risk over time 	Yes
OSP/I	<ul style="list-style-type: none"> Single factor – simple and likely to be consistently applied Outperformed the RM2000 Some evidence of significantly predicting IIOC reoffending May be useful with individuals with sexual offending to distinguish risk of IIOC recidivism 	<ul style="list-style-type: none"> Limited use in terms of informing approaches to risk management No formal inter-rater reliability evidence is available Need for replication with non-prison sample Not specifically for use with IIOC individuals Does not align with risk principle of effective practice 	Partial
ERIC SO	<ul style="list-style-type: none"> Detailed and structured approach including analysis Designed to be used with individuals with IIOC offending. 	<ul style="list-style-type: none"> No validation studies yet. Currently more of a framework for data collection 	No
Levels of Service	<ul style="list-style-type: none"> Draws from wide array of evidence and literature Based on Central 8 – a well established principle of effective practice Predicts general recidivism consistently Consistently good inter-rater reliability Evidence of moderate relationship between LSI-OR risks/ needs total score and CPORT Risk/needs score and initial risk level predictive of any recidivism Useful for informing case management and treatment targets 	<ul style="list-style-type: none"> Issues in predicting new IIOC offending, sexual offending, or violent offending. Overrides reduce predictive accuracy. Current findings with IIOC individuals are preliminary Can be used to predict general recidivism but not IIOC specific offending 	Partial

APPENDIX B: ASSESSMENT PROCESS MAP

Initial Assessment Process: Social Work

Initial Assessment Process: Social Work



APPENDIX C: PILOT SITE CRITERIA

If your local authority wishes to be considered as a pilot site then please complete the following section. Please provide evidence in the space provided of how your local authority is able to meet each of the criteria set out below.

Appropriate selection, training and support for staff

The framework being piloted should be applied by appropriately trained staff. As a new method it is also important that there are appropriate support mechanisms in place for staff involved in the pilot.

The following criteria therefore relate to this area, please indicate how you will meet each of these criteria –

1. Staff involved in the pilot should be Justice Social Workers who have been trained in the use of the LS/CMI (Scotland version) and have evidenced an adequate standard of assessment work. How will you identify and select appropriately qualified staff for the pilot.
2. How will line managers support those staff identified;
 - to take part in the pilot,
 - to complete the training on the CPORT and assessment framework,
 - with applying the framework to applicable cases,
 - to participate in inter-rater exercise following the training,
 - to participate in focus groups exploring the use of the framework.
3. Only individuals who have attended the necessary training (i.e. trained on CPORT/CASIC and use of the framework) will be considered competent to complete the assessment framework and eligible to participate in the pilot. Therefore, please indicate the number of staff to be involved in the pilot and who will require training.
4. Staff turnover is going to occur during the pilot. To support planning regarding training and support please indicate an estimate regarding rates of staff turnover across a 2 year period.
5. Throughout the pilot it is important that there is effective communication between Justice Social Workers, line management, senior management and the representation on the working group. Outline the communication and governance arrangements that will support the local authority to effectively engage in this project.

Stakeholder engagement

6. Please identify the stakeholders in your local authority area that you consider relevant in terms of needing to have an awareness of the pilot and how will you support them to have the relevant information?

Working Group representation

7. Each pilot site will need to identify representation for a working group that will convene throughout the duration of the pilot.

The responsibilities will involve;

- collating and recording feedback from practitioners applying the method,
- collating any live issues within their local authority and communicating them to the working group for consideration,
- identifying and disseminating solutions agreed at the working group,
- supporting practitioners in their authority in the use of the assessment framework.

Following identification of the pilot sites this group will initially meet monthly.

Please indicate who will be your local authority representative, their role and their commitment to the responsibilities listed above.

Data sharing & research

8. Please indicate how many internet offenders have been assessed by your service in the past 2 years, plus an estimate of how many might be assessed during the 2 years of the pilot.

9. Are you able to identify assessments previously completed on internet offenders within your local authority prior to the pilot?

10. Please provide an overview of how your current data capture mechanisms and case management system/s can support data sharing. This includes –

- a. information from the LS/CMI system report to support retrospective analysis,
- b. any previous assessments applied on internet offenders within your local authority,
- c. completed CPORT and assessment frameworks throughout the pilot.

Quality Assurance

11. What quality assurance mechanisms/methods are in place or will your service put in place to support the pilot (e.g. auditing of assessments, staff meetings, supervision).

12. How will learning identified during the course of the pilot and through evaluation be communicated effectively to staff?

RMA

**RISK
MANAGEMENT
AUTHORITY**

7 Thread Street
Paisley
PA1 1JR

0141 278 4478
info@rma.gov.scot

www.rma.scot