Original Research Article

Sexual Abuse
2024, Vol. 0(0) 1–18

DOI: 10.1177/10790632241290506
journals.sagepub.com/home/sax
Sage

Predictive Validity of Tools for Assessing Recidivism Risk in Men Convicted of Sex Offending: Static-99R, Static-2002R and BARR-2002R

Claire Ducro[1,2] and Thierry H. Pham[1,3]

Abstract

Recidivism risk assessment is crucial for effective case management of men convicted of sex offending. The use of empirical actuarial risk tools has become routine in the field. However, the development of actuarial risk scales for assessing general, violent and/or sexual recidivism in these men is ongoing: The Static-99 led to the Static-99R and the Static-2002R, and the BARR-2002R emerged to assess violent recidivism risk. A study was undertaken to evaluate and compare the inter-rater and predictive validity of the Static-99R, the Static-2002R, and the BARR-2002R in a sample of 328 men convicted of sex offending released from prison in French Belgium. When the instruments were considered integrally, the two versions of the Static—the Static-99R and the Static-2002R—proved better at predicting sexual recidivism and the BARR-2002R was better at predicting violent recidivism. And, the predictive and incremental predictive validity of the factor structure identified by Brouillette-Alarie et al. (2016) was examined. Results proved consistent in that the *Youthful stranger aggression* and *General criminality* factors were better at predicting general recidivism and violent non-sexual recidivism while the *Persistence/paraphilia* factor was better at predicting sexual recidivism.

[1]Centre de Recherche en Défense Sociale, Tournai, Belgium
[2]Univ. Lille, Villeneuve d'Ascq, France
[3]UMons, Université de Mons, Mons, Belgium

Corresponding Author:
Claire Ducro, Psychologie, Univ. Lille, CNRS, UMR 9193 – SCALab – Sciences cognitives et Sciences Affectives, Villeneuve d'Ascq, France.
Email: Claire.ducro@univ-lille.fr

Keywords
static-99R, static-2002R, BARR-2002R, men convicted of sex offending

Introduction

The development of actuarial instruments has made it possible to objectify how recidivism risk is predicted. Their performance in this field has compared favorably against non-structured professional judgment (Hanson & Morton-Bourgon, 2009; Harris et al., 2015). Over the years, a number of such instruments have been proposed for specific types of re-offending and target populations. Where men who offend sexually are concerned, the Static-99 and the Static-2002 (and their revised versions) were developed to predict risk of sexual re-offending (Hanson & Thornton, 2000; Harris et al., 2003; Helmus et al., 2012) and the BARR-2002 (and its revised version) was developed to predict risk of violent re-offending (Babchishin et al., 2013). Predictive validity refers to how well an instrument distinguishes recidivists from non-recidivists. One way of measuring this property is with the area under the curve (AUC; Mossman, 2013; Mossman & Somoza, 1991). To assess sexual recidivism risk in men who have offended sexually, the tools that are recommended are those that include items that deal specifically with sex offending, such as the Static-99R and the Static-2002R. Tools that do not examine sex offending are not recommended to assess recidivism risk (Brouillette-Alarie et al., 2016; Brouillette-Alarie & Lussier, 2018; Parent et al., 2011).

As part of the ongoing development of actuarial instruments, practitioners have been quick to want to determine which ones are better at predicting recidivism than others. This is understandable in that the assessment of recidivism risk is an issue of direct concern to decision-makers in the field of criminal justice and, because of this, of indirect concern to some health professionals.

It seemed worthwhile, then, to develop the predictive validity of the scores on the instruments mentioned above, discuss the divergent results regarding the usefulness of the age item in the instruments, and examine the predictive validity of the factors of these instruments, particularly the factor structure identified by Brouillette-Alarie et al. (2016).

Predictive Validity of Risk Instrument Total Scores

Regarding predictive validity, numerous studies have been undertaken to determine which risk assessment tool is best suited to a given purpose but results to date have been mixed and, therefore, no consensus has emerged in this regard (Reeves et al., 2018). According to some authors, though the difference is statistically modest, the Static-2002 is superior to the Static-99 in predicting sexual recidivism risk (Stalans et al., 2010). Other authors have found the Static-99/99R and the Static-2002/2002R to be comparable in this regard (Babchishin et al., 2012; Bengtson, 2008; Helmus & Hanson,

2007; Langton, Barbaree, Hansen, et al., 2007; Langton Barbaree, Seto, et al., 2007; Looman & Abracen, 2010). This said, it need be underscored that the convergent validity of these two instruments has been found to be very good (Kendall's Tau = .69; Jung et al., 2015). However, it is a fact that men who offend sexually also re-offend non-sexually (Ducro et al., 2020; Hanson & Bussiere, 1998; Menghini et al., 2005; Prentky et al., 1997). Consequently, this recidivism risk needs to be assessed as well. Violent recidivism risk is normally assessed using instruments such as the Static-99R or the Static-2002R, not to mention the SORAG (Sex Offenders Risk Appraisal Guide; Quinsey et al., 1995) and the VRAG-R (Violence Risk Appraisal Guide Revised; Rice et al., 2013). Babchishin et al. (2012) validated the predictive properties of certain items of the Static-2002R regarding violent recidivism and, particularly, those related to age and general criminality. It is on the basis of these findings and following a study of predictive validity and convergent validity with other tools measuring violent re-offending or antisocial behavior that these authors developed and validated the BARR-2002R. This instrument has received special attention because, as it is an integral part of the Static-2002R, the different types of recidivism can be evaluated by way of a single tool (Babchishin et al., 2012). A review of the literature to date supports the good predictive validity of the Static-99R, the Static-2002R and the BARR-2002R. Unlike the Static-99R, however, the Static-2002R and particularly the BARR-2002R have not been investigated as much in terms of predictive validity internationally.

Predictive Validity of Factors Identified in the Static-99R and Static-2002R

Practitioners who use risk assessment tools are more inclined to focus institutional or clinical case management on criminogenic needs, which are front and center when it comes to risk assessment. In other words, they place a greater emphasis on the main criminogenic factors. Risk assessment tools can also be useful in determining prison terms, whether to grant parole, or whether to impose certain types of supervision and other judicial measures. While static instruments predict recidivism, it is not clear whether evaluators can use static factors as treatment targets or change criteria (Brouillette-Alarie et al., 2016). Still, as these static factors constitute measures of observed behavior associated with recidivism, it should be possible to extract from them the latent psychological constructs responsible for the risk (Brouillette-Alarie et al., 2014). Factor analyses previously conducted on the Static-99 and the Static-99R have yielded various two- and three-factor structures (Brouillette-Alarie & Proulx, 2013). Based on these structures, underlying psychological mechanisms have been proposed to explain the factors. These mechanisms can then be targeted in the course of case management. One factor identified concerns sexual deviance probably measured by persistence of sexual offending, a second factor regards general criminality, which is interpreted more in terms of antisocial behavior with the notion of disregard for other people's rights and the notion of illegal impulses, and the third in the three-factor solution pertains to young age of person assessed (Brouillette-Alarie et al., 2014, 2016). Interpreting this last factor as a psychological phenomenon is not immediately obvious

but one way of doing so is to consider that younger sex offenders have not explored stable amorous relationships, are less likely to have children, and are more likely to victimize strangers perhaps with greater coercion. The aim of the researchers who carried out these factor analyses was to explore more in depth the constructs underlying the two versions of the Static, that is, the Static-99R and the Static-2002R. These two tools are widely used, easy and fast to use given that they require rating risk based on an examination of a person's criminal record, and used in numerous countries around the world. Analyzing latent structure affords a fresh look on the static assessment of recidivism risk. For the purposes of case management planning, dynamic recidivism risk assessment tools such as the Stable-2007 (Hanson et al., 2007) are used in a limited and complementary manner on account of institutional constraints. More specifically, these tools take a relatively long time to administer and must be re-administered periodically over time. However, they have demonstrated good convergent validity with the Static-99R and the Static-2002R. According to Brouillette-Alarie et al. (2014), "actuarial tools composed of static risk factors could inform practitioners, albeit imperfectly, regarding the treatment needs of people who offend sexually" [free translation].

Regarding the predictive validity of the factor structure put forth by Brouillette-Alarie (2016), the three factors identified are significantly and positively associated with sexual recidivism. General recidivism and non-sexual violent recidivism are more significantly and positively associated with the *Youthful stranger aggression* and *General criminality* factors. Validating this factor structure allows understanding why some tools are more predictive of certain types of recidivism. It also helps understanding the contradictory results between tools.

Study Objectives

Against this background, we conducted a study to evaluate the predictive validity of the Static-99R, the Static-2002R and the BARR-2002R among men convicted of sex offending and later released from prison in French Belgium, over two follow-up periods, namely, five years, which was similar to periods used in the international literature, and nearly ten years.

Finally, further to the work of Brouillette-Alarie (Brouillette-Alarie et al., 2016) aimed at identifying the underlying factors that explain the psychological processes at play, we evaluated the predictive and incremental predictive validity of these factors.

Method

Study Group

Our study examined 328 adult men convicted of sex offending in French Belgium and who received treatment or guidance from a Specialized Health Team (SHT) in 2001 and/or 2002 or treatment from 2009 to 2015. This study population was composed

of the populations of two earlier studies (not published). The two cohorts did not differ significantly on the variables measured, i.e., in terms of level of risk assessed with the different instruments used in our study. Of the men in the study population, 23.46% had offended exclusively against victims 14 or over, 60.80% against victims under 14, and 15.74% against both victims 14 or over and victims under 14. It should be noted that the age criterion is widely used in the international literature and mentioned in the official criminal records of sex offenders in Belgium.

Mean age of study group was 43.21 years, $SD = 11.68$. Age at release ranged from 18.27 to 79.22 years.

Post-release period (i.e., period from release to recidivism or, in the case of non-recidivists, to research end date) ranged from 0 to 22.49 years, with a mean duration of 9.51 years, $SD = 5.61$.

Instruments

Static-99R. The Static-99R (Hanson & Thornton, 2000; Helmus et al., 2012) is a static, actuarial instrument for assessing sexual recidivism risk in adult men who offended sexually. It comprises 10 items regarding the characteristics of past and present sex and non-sex offenses committed and the specificity of the victims of sex offenses. Items regarding age of offender and whether offender ever lived with an intimate partner are also taken into account. In its revised version, the Static-99R allows determining risk level based on a total score with a possible range of −3 to 12, which breaks down into five risk categories. It should be noted that the main difference between the Static-99 and its revised version regards how the age item is scored and the scoring criteria used (Reeves et al., 2018).

The instrument has demonstrated high inter-rater reliability, with intraclass correlation coefficients (ICCs) ranging from .84 to .95 across 11 studies reviewed by Phoenix et al., in 2016. The predictive validity reported in the international literature as measured by the AUC has been moderate to high. AUC values have averaged .72 for general recidivism, .68 for sexual recidivism, .69 for violent sexual recidivism, and .70 for violent non-sexual recidivism (Babchishin et al., 2012, 2016).

Static-2002R. The Static-2002R (Helmus et al., 2012) was developed to assess sexual recidivism risk in adult men convicted of sex offending. This is a revised version of the Static-2002 (Phenix et al., 2008). It is composed of 14 items covering five factors: One item regards age at time of offense, three items regard persistence of sex offending, three items regard deviant sexual interests, two items regard relationship to victims, and five items regard general criminality. Items are scored based on information gleaned from the offender's criminal record. These scores yield a total score with a possible range of 0 to 14, which breaks down into five sexual recidivism categories. In terms of psychometric properties, the Static-2002 has demonstrated very good inter-rater agreement, obtaining an ICC of .98, and high predictive validity, with AUC values of .76 for general recidivism, .70 for violent recidivism, and .76 for sexual recidivism

(Helmus & Hanson, 2007). The creators of the original Static-2002 made one key change to the instrument when they revised it: They assigned a different weight to the age item for young men. The Static-2002R total score has a possible range of −2 to 13. This revised version obtained higher AUC values: .76 for general recidivism, .69 for violent recidivism (including violent sexual recidivism), .74 for violent non-sexual recidivism, and .77 for non-violent non-sexual recidivism.

Brief Assessment of recidivism risk (BARR-2002R; Babchishin et al., 2013). Based on existing recidivism risk assessment tools and particularly on the Static-2002R, the creators of the Static-2002 removed the sexual items from the instrument and sought to validate whether the remaining items were predictive of non-sexual recidivism risk. This exercise led to the development of the BARR-2002R, an instrument for assessing violent recidivism risk in men who offended sexually. In other words, this instrument focuses on notions of general criminality. It has demonstrated high predictive validity, obtaining AUC values of .77 for general recidivism, .72 for violent recidivism, and .74 for violent non-sexual recidivism (Babchishin et al., 2016). Jung et al. (2015) reported rather similar values: .72 for general recidivism, .74 for violent recidivism, and .66 for sexual recidivism.

Procedure

We examined the criminal records of the men in the study group to glean date of release, date of recidivism (if any), type of recidivism (if any), and information to assess recidivism risk level using the Static-99R, the Static-2002R and the BARR-2002R. Records were consulted by two evaluators at the Ministry of Justice in order to be able to access all the documents required to complete the risk assessment tools and to determine, in double blind, presence of recidivism and type. The documents examined included statements of the facts (offenders and victims), criminal records, prison admission records, psychiatric and psychological evaluations, and psychosocial evaluation reports drafted by correctional services.

Four types of recidivism were defined based on criminal convictions by final judgment: general recidivism corresponded to any new conviction regardless of offense; sexual recidivism corresponded to any new conviction for a sex offense; violent non-sexual recidivism corresponded to any new conviction for a violent non-sexual offense, such as non-sexual physical assault; and non-violent non-sexual recidivism corresponded to any new conviction for a non-violent non-sexual offense, such as theft.

Two raters examined the records of the same 75 men selected at random for the purpose of determining inter-rater agreement on the instruments. Agreement was very good as evidenced by the ICCs obtained: .97 for the Static-99R total score, .91 for the BARR-2002R total score, and .93 for the Static-2002R total score. Inter-rater agreement coefficients obtained for the four Static-2002R factors were .87 for Persistence, .95 for Deviance, .91 for Relationship with victims, and .88 for General

criminality. Once inter-rater agreement had been measured, the two raters split the remaining records at random.

Completing the different instruments allowed us first to obtain scores for each of these and for the factors of the Static-2002R. Second, in order to run analyses and fulfill our study's objectives, we also calculated the score for each factor of the structure proposed by Brouillette-Alarie and colleagues (2016). This structure is based on the factor analysis of the combined Static-99R and Static-2002R. Three factors were identified. *Persistence/paraphilia* is the sum, from 0 to 7, of prior sentencing occasions for sexual offenses (Static-2002R), rate of sexual offending (Static-2002R), any sentencing occasion for non-contact sex offenses (Static-2002R), any male victim (Static-2002R), young, and unrelated victims (Static-2002R). *Youthful stranger aggression* is the sum, from −2 to 7, of ever lived with lover (Static-99R), age at release (Static-2002R), any unrelated victim (Static-2002R), any stranger victim (Static-2002R), index non-sexual violence (Static-99R), any juvenile arrest for a sexual offense, and convicted as an adult for a separate sexual offence (Static-2002R). *General criminality* is the sum, from 0 to 6, of any prior involvement with the criminal justice system (Static-2002R), prior sentencing occasions for anything (Static-2002R), any prior non-sexual violence sentencing occasion (Static-2002R), any community supervision violation (Static-2002R), and years free prior to index sex offense (Static-2002R).

Data Analyses

Convergent validity was determined by calculating *r* between the Static-99R, the Static-2002R and the BARR-2002R. Convergent validity will be discussed according to Cohen's criteria: .10 is small convergent validity, .30 is medium convergent validity and. 50 is large convergent validity. The predictive validity of the instruments was evaluated using ROC analysis and rated according to the following guidelines proposed by Cohen (Cohen, 1992; Rice & Harris, 2005): AUC values above .714 correspond to a large effect size, values from .639 to .714 correspond to a moderate effect size, and values below .639 correspond to a small effect size. The ROC area (AUC) represents the probability that a randomly selected recidivist will obtain a higher risk score than will a randomly selected non-recidivist (Rice & Harris, 2005).

Predictive validity was calculated first for the instrument total scores, second for the factor structure identified from the Static-99R and Static-2002R. Validity was calculated for three follow-up periods: 0–5 years, fixed 5 or more years, and a longer period with a mean of 9.57 years.

Cox regressions were run for the factor structure identified for the Static-99R and Static-2002R. These are similar to logistic regressions in that both methods predict outcomes on the basis of a set of variables, in our case, the three factors identified. However, Cox regressions also take account of the time elapsed before the outcome. The relevant coefficients in these regressions are the Wald statistical significance value and the hazard ratio. The Wald value indicates whether an independent variable actually

predicts the outcome. The hazard ratio gauges the strength of the association between each independent variable and the outcome (the dependent variable). A hazard ratio greater than 1 means that an increase in the independent variable increases the risk of the event (dependent variable) occurring, whereas a hazard ratio less than 1 indicates that an increase in the independent variable reduces that risk.

Results

Convergent Validity

Convergent validity (Table 1) proved high between the Static-99R and the Static-2002R total scores, between the Static-99R and the BARR-2002R total scores, and between the Static-2002R and the BARR-2002R total scores.

Regarding the instrument total scores, with or without the age item, that is, regardless of the calculation method used, the instrument factors generally presented large convergent validity, except for the *Sexual Deviance* factor, which presented a medium convergent validity. As for the factor structure identified by Brouillette-Alarie et al. (2016), generally speaking, predictive validity was good for the instrument total scores, with some nuances for the Static-99R and the BARR-2002R. However, what was interesting was that the convergent validity between the factors was not as strong. Indeed, according to Cohen's criteria, it was medium to small at best when it managed to be statistically significant.

Predictive Validity

Before presenting the results regarding instrument predictive validity, it seems important to reiterate the study population's recidivism rates. Over the fixed 5-year follow-up period (N = 229), the most common type of recidivism was sexual recidivism (8.7%), followed by non-violent non-sexual recidivism (2.2%); general recidivism totaled 10.5%. Lastly, when recidivism rates were calculated for the entire study population over a mean period of 9.22 years (N = 334), sexual recidivism was the most common type of recidivism, reaching 12.9%, followed by non-violent non-sexual recidivism (11.1%) and violent non-sexual recidivism (4.8%); general recidivism was 25.1%.

Generally speaking, the three instruments predicted the types of recidivism considered with a moderate to large effect size (Tables 2 and 3).

General recidivism was predicted primarily by all the instrument total scores with a moderate effect size, but also by the *General criminality* factor of the Static-2002R and the *Youthful stranger aggression* and *General criminality* factors of the structure identified by Brouillette-Alarie (2016). When the predictive validity of these instruments for general recidivism was considered over a longer follow-up period, effect sizes diminished slightly but remained statistically significant. Still, when a fixed follow-up period of five years was considered, only the Static-99R and the

Table 1. Convergent Validity and Mean Scores of Instruments and Their Factor Structure.

	Pearson's *r*										
	1	2	3	4	5	6	7	8	9	M	*SD*
1. Static-99 R: Total score	1									2.05	2.31
2. Static-2002R: Total score	**.880****									3.33	2.51
3. Factor: Persistence of sexual offending	**.400****	**.639****								0.42	0.84
4. Factor: Deviant sexual offending	**.279****	**.437****	**.258****							0.63	0.90
5. Factor: Relationship to victims	**.688****	**.627****	**.234****	**.168****						0.68	0.74
6. Factor: General criminality	**.499****	**.629****	**.431****	.035	**.268****					1.07	1.04
7. BARR-2002R: Total score	**.709****	**.737****	**.310****	−.053	**.352****	**.834****				2.30	2.12
Brouillette-Alarie et al. Structure											
8. Persistence: paraphilia	**.436****	**.672****	**.780****	.769**	**.242****	**.300****	**.168****			1.09	1.43
9. Youthful stranger aggression	**.823****	**.631****	.070	−.026	**.668****	**.174****	**.555****	.020		1.80	1.87
10. General criminality	**.482****	**.591****	**.395****	.009	**.249****	**.968****	**.849****	**.266****	**.163****	1.82	1.82

**$p < .001$; *$p < .05$. Bold:significative results.

Table 2. Predictive Validity of the Static-99R, the Static-2002R, and the BARR-2002R (Follow-up: M = 9.57 years).

	AUC (95% CI)			
	General recidivism	Sexual recidivism	Violent non-sexual recidivism	Non-violent non-sexual recidivism
	Follow-up: 0–22 years, *M* = 9.57, *SD* = 5.61			
Static-99R: Total score	**.70** (.64–.77)**	**.68** (.60–.77)**	**.74** (.63–.86)**	**.69** (.57–.84)**
Static-2002R: Total score	**.69** (.62–.75)**	**.68** (.59–.77)**	**.68* (.57–.80))**	**.67** (.58–.76)**
Factor: Persistence of sexual offending	.56 (.48–.64)	.58 (.48–.69)	.45 (.31–.60)	.57 (.46–.67)
Factor: Deviant sexual offending	.56 (.48–.63)	**.67** (.58–.76)**	.53 (.38–.68)	.47 (.37–.58)
Factor: Relationship to victims	**.58* (.51–.66)**	**.62* (.53–.72)**	.59 (.44–.74)	.52 (.42–.63)
Factor: General criminality	**.62** (.55–.70)**	.53 (.43–.64)	.65 (.52–.78)	**.70** (.62–.79)**
BARR-2002R: Total score	**.70** (.64–.77)**	.60 (.51–.69)	**.74** (.63–.86)**	**.75** (.67–.84)**
Brouillette-Alarie et al. structure				
Persistence: paraphilia	**.58* (.50–.65)**	**.68** (.59–.77)**	.48 (.329–.623)	.53 (.428–.641)
Youthful stranger aggression	**.61** (.60–.74)**	**.65** (.56–.73)**	**.73** (.62–.85)**	**.63*(.53–.72)**
General criminality	**.62** (.54–.69)**	.52 (.41–.63)	.65 (.51–.78)	**.70** (.61–.79)**

**$p < .001$; *$p < .05$. Bold:significative results.

Static-2002R and, more specifically, their *Deviant sexual offending* and *Youthful stranger aggression* factors, proved significant predictors of general recidivism.

Sexual recidivism over a long follow-up period of five years was predicted with a small effect size by the Static-99R, the Static-2002R, and the *Persistence* factor of the structure identified by Brouillette-Alarie et al. (2016). It was predicted with a more moderate effect size by the Static-2002R *Sexual deviance* factor, the BARR-2002R total score, and the *Youthful stranger aggression* factor of the structure identified by Brouillette-Alarie and colleagues (2016). However, only the *Youthful stranger*

Table 3. Predictive Validity of the Static-99R, the Static-2002R, and the BARR-2002R (Follow-up Fixed at 5 years).

	AUC (95% CI)			
	General recidivism	Sexual recidivism	Violent non-sexual recidivism	Non-violent non-sexual recidivism
Static-99R: Total score	**.68** (.47–.73)**	.64 (.51–.77)	.59 (.49–.69)	**.90** (.80–1.00)**
Static-2002R: Total score	**.67* (.55–.79)**	.63 (.50–.76)	.70 (.60–.79)	**.89** (.81–.97)**
Factor: Persistence of sexual offending	.49 (.35–.62)	.47 (.33–.61)	.41 (.00–.88)	.68 (.37–.99)
Factor: Deviant sexual offending	**.64* (.51–.77)**	.62(.49–.76)	.88 (.80–.97)	.74 (.47–1.00)
Factor: Relationship to victims	.62 (.49–.75)	.62(.48–.76)	.68 (.46–.89)	.64 (.37–.90)
Factor: General criminality	.47 (.32–.62)	.46 (.29–.62)	.23 (.00–.50)	.62 (.37–.87)
BARR-2002R: Total score)	.60 (.47–.73)	.57 (.43–.72)	.38 (.26–.49)	.76 (.57–.95)
Brouillette-Alarie et al. structure				
Persistence: paraphilia	.60 (.48–.73)	.59 (.46–.71)	.78 (.68–.89)	.76 (.45–1.00)
Youthful stranger aggression	**.69** (.57–.80)**	**.66* (.54–.79)**	.64 (.52–.77)	.79 (.57–1.00)
General criminality	.46 (.31–.61)	.45 (.28–.61)	.22 (.00–.48)	.60 (.33–.86)

**$p < .001$; *$p < .05$. Bold:significative results.

aggression factor proved a predictor over the follow-up period fixed at 5 years. Tables 3 and 4.

For the violent or non-violent non-sexual recidivism, the AUC values varied very little when a longer follow-up period was considered. Still, over a fixed period of 5 years, no instrument predicted violent non-sexual recidivism, although the Static-99R and the Static-2002R were predictive of non-violent non-sexual recidivism. However, these results must be contextualized, given that the recidivism rates were relatively low (respectively, 0.4% and 2.2%).

Table 4. Incremental validity: Cox Regression.

	Persistence: Paraphilia				Youthful stranger aggression				General criminality			
	β (S or NS)	Wald	HR	95% CI	β	Wald	HR	95% CI	β	Wald	HR	95% CI
General recidivism	.10 (NS)	1.57	1.11	0.94–1.30	**.27 (S)**	**19.47**	**1.30**	**1.16–1.47**	**.18 (S)**	**7.91**	**1.20**	**1.06–1.35**
Sexual recidivism	**.40 (S)**	**13.75**	**1.50**	**1.21–1.85**	**.32 (S)**	**13.26**	**1.37**	**1.16–1.63**	.00 (NS)	0.00	0.98	0.83–1.21
Violent non-sexual recidivism	−.17 (NS)	0.62	0.85	0.56–1.28	**.39 (S)**	**7.55**	**1.47**	**1.13–1.93**	**.29 (S)**	**3.94**	**1.30**	**1.00–1.77**
Non-violent non-sexual recidivism	.01 (NS)	0.01	1.01	0.80–1.27	**.19 (S)**	**4.46**	**1.21**	**1.01–1.43**	**.35 (S)**	**14.42**	**1.42**	**1.19–1.71**

Bold:significative results.

Incremental Predictive Validity

The predictive validity of each factor (*Persistence, Youthful stranger aggression* and *General criminality)* was assessed by way of Cox regression analysis with recidivism (general, sexual, violent non-sexual, and non-violent non-sexual recidivism) as the outcome. One model was tested per type of recidivism (Table 4). General recidivism was significantly predicted by the *Youthful stranger aggression* and *General criminality* factors, which also predicted violent non-sexual recidivism and non-violent non-sexual recidivism. Sexual recidivism, however, was predicted by the *Persistence* and *Youthful stranger aggression* factors. In other words, the *Youthful stranger aggression* factor predicted all four types of recidivism in conjunction with either the *Persistence* factor for sexual recidivism or the *General criminality* factor for general recidivism, violent non-sexual recidivism, and non-violent non-sexual recidivism.

Discussion

Static variables and historical criminal factors are predictors of recidivism in that they reveal risk propensity related to antisocial orientation and sexual interests (Mann et al., 2010). Identifying the constructs responsible for recidivism is essential in order to help the health and criminal justice systems make decisions regarding institutional and therapeutic management paths. The Static-99 was originally developed to predict sexual and violent recidivism (Harris et al., 2003). However, research on this instrument and its revised version and on the Static-2002 and its revised version has suggested that general and violent recidivism might be better predicted by only a portion of the Static-2002R. These findings led to the creation of the BARR-2002R, which is composed of the Static-2002R's *age at release* item and its items associated with the *General criminality* factor (Babchishin et al., 2013).

Convergent Validity of Instruments

Our analysis of the convergent validity of the instruments revealed a very high convergence between the instruments, which is in line with past research on the design and validation of these three instruments (Babchishin et al., 2016; Helmus et al., 2012). It should be noted that the three have a number of items in common and that their development was geared to improving their predictive validity. Babchishin et al. (2012), for instance, pointed out that both versions of the Static should be used in order to estimate the recidivism risk posed by an offender. However, our analysis of the convergent validity of the factor structure of the instruments shows this validity to be from moderate to weak at best when it managed to reach statistical significance. This might be explained by the fact that the factors identified evaluate different concepts, as the work by Brouillette-Alarie and colleagues (2016) has tended to demonstrate. This said, the different factors identified present good convergent validity with the total scores on the different instruments considered in this study.

Given the evolution of recidivism risk assessment instruments and the growing interest in risk assessment among practitioners (Singh et al., 2016), it seems necessary to evaluate the predictive properties of the Static-99R, the Static-2002R and the BARR-2002R internationally as well. We examined the predictive validity of the instruments not only integrally but also with the age item omitted in light of the controversy in the international literature in this regard. Moreover, we examined the predictive validity of the factor structure identified by Brouillette-Alarie and colleagues in 2016.

Predictive Validity

It is clearly established in the international literature that use of the Static-99R or the Static-2002R is appropriate for assessing sexual recidivism and that use of the BARR-2002R is appropriate for assessing non-sexual recidivism. The primary aim of our study was to evaluate the predictive validity of these instruments in a French-speaking Belgian population. Our analyses of predictive validity show that, by and large, the tools predict the different types of recidivism with a significant effect size. It should be noted that the magnitude of this validity tended to diminish somewhat for very long follow-up periods, and that it tended not to be statistically significant for fixed follow-up periods of 5 years or more.

In sum, the three instruments predict general recidivism similarly, which is to say moderately well. However, given the results for the other types of recidivism, we would recommend using either version of the Static to predict sexual recidivism and the BARR-2002R to predict violent non-sexual recidivism and non-violent non-sexual recidivism, which is in keeping with the genesis of the BARR-2002R. This interesting finding nonetheless needs to be replicated in other scientific studies.

If we wish to gain a better understanding of the usefulness of certain items or combinations thereof, subjecting the instruments to factor analyses and then evaluating the predictive validity of the structures to emerge might afford a fresh perspective on risk assessment. The Static-99R has already been subjected to such analyses in the aim of identifying the psychological processes underlying its factors. It would be interesting to carry out this sort of study in conjunction with an investigation of the predictive validity of the factors with the other static instruments for assessing recidivism risk. This sort of research needs to be pushed further and, at the very least, repeated on a larger scale. This would contribute to advance knowledge that could help practitioners interpret their risk assessment results. In French Europe, in the course of training on the use of risk assessment instruments, we are often faced with the following questions: *What type of instrument should be used? Which items are most predictive? If you do not wish to express recidivism risk in terms of a score, category or likelihood, how can you organize your assessment report by focusing on the items predictive of risk?* These questions might seem odd to the eyes of practitioners in the habit of reporting numbers as recommended by the creators of the various tools. In practice, however, it turns out that the decision-makers for whom these reports are prepared, who may not be used to such rating systems, prefer to understand why men who have offended sexually pose a

recidivism risk, instead of merely being fed a score. The analyses that we carried out for the purposes of our study and efforts such as by Brouillette-Alarie et al. (2016) to structure risk by factors could help practitioners draft their reports more clearly and effectively. In this regard, we could highlight the factors of instruments with significant predictive properties and the likelihood of re-offending associated with these factors, as is presently proposed for the instrument total scores. The results of our study show, in accordance with what was reported by Brouillette-Alarie et al. (2016), that the *Youthful stranger aggression* and *General criminality* factors are more predictive of general recidivism and violent non-sexual recidivism and that the factor more specific to persistence of sexual offending or paraphilia is predictive of sexual recidivism. This underscores the importance of being mindful of the heterogeneousness of sex offenders in terms not only of their potential recidivism risk but also of how they are managed, whether clinically or institutionally.

In closing, though the matter warrants more in-depth research, our study allows us to recommend that practitioners use the BARR-2002R rather than the Static-99R or the Static-2002R total score to predict general recidivism and violent recidivism in men who have offended sexually. The BARR-2002R can be used more specifically to evaluate antisocial propensities. Men who have offended sexually and who score high on the BARR-2002R may have special management needs in connection with these dimensions, in addition to those in connection with sex-related issues. In this light, choosing to use one instrument or another depending on the purpose being pursued, analyzing the predictive factors of these instruments and giving some thought to their underlying psychological processes would allow evaluators to give a stronger clinical meaning to their evaluations and to the implications that these have.

Limitations and Outlook

Our study is not without limitations. For one, population size and recidivism rates might seem smaller compared with other studies. However, our study focused exclusively on a Belgian population of French-speaking inmates. It did not include French- or Dutch-speaking Belgian forensic psychiatric inpatients or Dutch-speaking inmates. Consequently, our sample was not representative of the Belgian population of persons who have offended sexually. Though the size of the sample is small by international standards, it is nevertheless encouraging. Future studies should seek to develop this line of research further and to carry out comparative analyses of predictive validity by sex offender type and, similarly, exploratory or confirmatory factor analyses with a sufficiently large sample.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This study was supported by the Ministry of Health, Social Affairs and Equal Opportunity of the Walloon Region of Belgium.

ORCID iD

Claire Ducro https://orcid.org/0000-0002-4488-3496

Data Availability Statement

The authors take responsibility for the integrity of the data and for the accuracy of the data analyses, and have made every effort to avoid inflating statistically significant results.

References

Babchishin, K. M., Hanson, R. K., & Blais, J. (2013). User guide for the brief assessment for recidivism risk – 2002R (BARR-2002R). https://www.static99.org/pdfdocs/Static2002-BARR-2002REvaluatorsHandbook-2013-12-03.pdf

Babchishin, K. M., Hanson, R. K., & Blais, J. (2016). Less is more: Using Static-2002R subscales to predict violent and general recidivism among sexual offenders. *Sexual Abuse*, *28*(3), 187–217. https://doi.org/10.1177/1079063215569544

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

Bengtson, S. (2008). Is newer better? A cross-validation of the static-2002 and the risk matrix 2000 in a Danish sample of sexual offenders. *Psychology, Crime and Law*, *14*(2), 85–106. https://doi.org/10.1080/10683160701483104

Brouillette-Alarie, S., Babchishin, K. M., Hanson, R. K., & Helmus, L.-M. (2016). Latent constructs of the static-99r and static-2002R: A three-factor solution. *Assessment*, *23*(1), 96–111. https://doi.org/10.1177/1073191114568114

Brouillette-Alarie, S., Hanson, R. K., Babchishin, K. M., & Benbouriche, M. (2014). De la prédiction à la compréhension : recension des dimensions psychologiques de la Statique-99 [From prediction to understanding: A literature review of the psychological dimensions of the Static-99]. *Pratiques Psychologiques*, *20*, 1–19. https://doi.org/10.1016/j.prps.2013.12.001

Brouillette-Alarie, S., & Proulx, J. (2013). Predictive validity of the Static-99R and its dimensions. *Journal of Sexual Aggression*, *19*(3), 311–328. https://doi.org/10.1080/13552600.2012.747630

Brouillette-Alarie, S., Proulx, J., & Hanson, R. K. (2018). Three central dimensions of sexual recidivism risk: Understanding the latent constructs of Static-99R and Static-2002R. *Sexual Abuse*, *30*(6), 676–704. https://doi.org/10.1177/1079063217691965

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155–159. https://doi.org/10.1037//0033-2909.112.1.155

Ducro, C. D., Telle, E., & Pham, T. (2020). Recidivism rates of sex offenders released from French Belgian judicial system. *Acta Psychiatrica Belgica*, *120*(1), 43–49.

Hanson, R. K., & Bussiere, M. T. (1998). Predicting relapse: A meta-analysis of sexual offender recidivism studies. *Journal of Consulting and Clinical Psychology*, *66*(2), 348–362. https://doi.org/10.1037/0022-006X.66.2.348

Hanson, R. K., Harris, A. J. R., Scott, T.-L., & Helmus, L. (2007). *Assessing the risk of sexual offenders on community supervision: The Dynamic Supervision Project (User report, Corrections research)*. Public Safety Canada. https://www.publicsafety.gc.ca/cnt/rsrcs/pblctns/ssssng-rsk-sxl-ffndrs/index-eng.aspx

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

Hanson, R. K., & Thornton, D. (2000). Improving risk assessments for sex offenders: A comparison of three actuarial scales. *Law and Human Behavior*, *24*(1), 119–136. https://doi.org/10.1023/A:1005482921333

Harris, A., Phenix, A., Hanson, R. K., & Thornton, D. (2003). *Static 99: Coding rules revised 2003*. Solicitor General of Canada. https://www.static99.org/pdfdocs/static-99-coding-rules_e.pdf

Harris, G. T., Rice, M. E., Quinsey, V. L., & Cormier, C. A. (2015). *Violent offenders: Appraising and managing risk* (3rd edition). American Psychological Association.

Helmus, L., Thornton, D., Hanson, R. K., & Babchishin, K. M. (2012). Improving the predictive accuracy of Static-99 and Static-2002 with older sex offenders: Revised age weights. *Sexual Abuse*, *24*(1), 64–101. https://doi.org/10.1177/1079063211409951

Helmus, L. M., & Hanson, R. K. (2007). Predictive validity of the Static-99 and Static-2002 for sex offenders on community supervision. *Sexual Offender Treatment*, *2*(2), 1–14. https://doi.org/10.1177/1079063217712216

Jung, S., Ennis, L., Hermann, C. A., Pham, A. T., Choy, A. L., Corabian, G., & Hook, T. (2015). An evaluation of the reliability, construct validity, and factor structure of the Static-2002R. *International Journal of Offender Therapy and Comparative Criminology*, *61*(4), 464–487. https://doi.org/10.1177/0306624X15595228

Langton, C. M., Barbaree, H. E., Hansen, K. T., Harkins, L., & Peacock, E. J. (2007a). Reliability and validity of the Static-2002 among adult sexual offenders with reference to treatment status. *Criminal Justice and Behavior*, *34*(5), 616–640. https://doi.org/10.1177/0093854806296851

Langton, C. M., Barbaree, H. E., Seto, M. C., Peacock, E. J., Harkins, L., & Hansen, K. T. (2007b). Actuarial assessment of risk for reoffense among adult sex offenders: Evaluating the predictive accurary of the Static-2002 and five other instruments. *Criminal Justice and Behavior*, *34*(1), 37–59. https://doi.org/10.1177/0093854806291157

Looman, J., & Abracen, J. (2010). Comparison of measures of risk for recidivism in sexual offenders. *Journal of Interpersonal Violence*, *25*(5), 791–807. https://doi.org/10.1177/0886260509336961

Mann, R. E., Hanson, R. K., & Thornton, D. (2010). Assessing risk for sexual recidivism: Some proposals on the nature of psychologically meaningful risk factors. *Sexual Abuse*, *22*(2), 191–217. https://doi.org/10.1177/1079063210366039

Menghini, M., Ducro, C., & Pham, T. (2005). Evaluation de la récidive auprès d'agresseurs sexuels issus d'un hôpital sécuritaire en Belgique francophone [Assessing recidivism among sex offenders released from a secure forensic hospital in French Belgium]. *Forensic*, 39–48. [Special issue].

Mossman, D. (2013). Evaluating risk assessments using receiver operating characteristic analysis: Rationale, advantages, insights, and limitations. *Behavioral Sciences & the Law*, *31*(1), 23–39. https://doi.org/10.1002/bsl.2050

Mossman, D., & Somoza, E. (1991). ROC curves, test accuracy, and the description of diagnostic tests. *The Journal of Neuropsychiatry and Clinical Neurosciences*, *3*(3), 330–333. https://doi.org/10.1176/jnp.3.3.330

Parent, G., Guay, J.-P., & Knight, R. A. (2011). An assessment of long-term risk of recidivism by adult sex offenders: One size doesn't fIt all. *Criminal Justice and Behavior*, *38*(2), 188–209. https://doi.org/10.1177/0093854810388238

Phenix, A., Doren, D., Helmus, L., Hanson, R. K., & Thornton, D. (2008). *Coding rules for static-2002*. Public Safety Canada. https://www.publicsafety.gc.ca/cnt/rsrcs/pblctns/sttc-2002/index-en.aspx

Prentky, R. A., Lee, A. F. S., Knight, R. A., & Cerce, D. (1997). Recidivism rates among child molesters and rapists: A methodological analysis. *Law and Human Behavior*, *21*(6), 635–659. https://doi.org/10.1023/A:1024860714738

Quinsey, V. L., Rice, M. E., & Harris, G. T. (1995). Actuarial prediction of sexual recidivism. *Journal of Interpersonal Violence*, *10*(1), 85–105. https://doi.org/10.1177/088626095010001006

Reeves, S. G., Ogloff, J. R., & Simmons, M. (2018). The predictive validity of the Static-99, Static-99R, and Static-2002R: Which one to use? *Sexual Abuse*, *30*(8), 887–907. https://doi.org/10.1177/1079063217712216

Rice, M. E., & Harris, G. T. (2005). Comparing effect sizes in follow-up studies: ROC area, Cohen's d, and r. *Law and Human Behavior*, *29*(5), 615–620. https://doi.org/10.1007/s10979-005-6832-7

Rice, M. E., Harris, G. T., & Lang, C. (2013). Validation of and revision to the VRAG and SORAG: The violence risk Appraisal guide — revised (VRAG-R). *Psychological Assessment*, *25*(3), 951–965. https://doi.org/10.1037/a0032878

Singh, J. P., Bjorkly, S., & Fazel, S. (2016). *International perspectives on violence risk assessment*. Oxford University Press.

Stalans, L. J., Hacker, R., & Talbot, M. E. (2010). Comparing nonviolent, other-violent, and domestic batterer sex offenders: Predictive accuracy of risk assessments on sexual recidivism. *Criminal Justice and Behavior*, *37*(5), 613–628. https://doi.org/10.1177/0093854810363794