Improving the Predictive Accuracy of Static-99 and Static-2002 With Older Sex Offenders : Revised Age Weights

Leslie Helmus, David Thornton, R. Karl Hanson and Kelly M. Babchishin

Sex Abuse 2012 24: 64 originally published online 15 August 2011
DOI: 10.1177/1079063211409951

Sexual Abuse: A Journal of Research and Treatment
24(1) 64–101

DOI: 10.1177/1079063211409951
http://sax.sagepub.com
SAGE

Improving the Predictive Accuracy of Static-99 and Static-2002 With Older Sex Offenders: Revised Age Weights

Leslie Helmus[1,2], David Thornton[3], R. Karl Hanson[1], and Kelly M. Babchishin[1,2]

Abstract

Actuarial risk assessment scales and their associated recidivism estimates are generally developed on samples of offenders whose average age is well below 50 years. Criminal behavior of all types declines with age; consequently, actuarial scales tend to overestimate recidivism for older offenders. The current study aimed to develop a revised scoring system for two risk assessment tools (Static-99 and Static-2002) that would more accurately describe older offenders' risk of recidivism. Using data from 8,390 sex offenders derived from 24 separate samples, age was found to add incremental predictive validity to both Static-99 and Static-2002. After creating new age weights, the resulting instruments (Static-99R and Static-2002R) had only slightly higher relative predictive accuracy. The absolute recidivism estimates, however, provided a substantially better fit for older offenders than the recidivism estimates from the original scales. We encourage evaluators to adopt the revised scales with the new age weights.

Keywords

risk assessment, recidivism, prediction, age, sex offenders

[1]Public Safety Canada, Ottawa, Ontario, Canada
[2]Carleton University, Ottawa, Ontario, Canada
[3]Sand Ridge Secure Treatment Center, Mauston, WI, USA

Corresponding Author:
Leslie Helmus, 10th floor, 340 Laurier Avenue West, Ottawa, Ontario, Canada, K1A 0P8
Email: Leslie.helmus@ps.gc.ca

Risk assessment is a key activity in criminal justice systems, with profound consequences for public safety and the offender. Actuarial risk assessment scales include explicit rules to combine prespecified items into total scores and provide empirically derived estimates of recidivism probability linked to each total score (Dawes, Faust, & Meehl, 1989). Actuarial scales are increasingly being used, particularly for sex offenders (Archer, Buffington-Vollum, Stredny, & Handel, 2006; Jackson & Hess, 2007; McGrath, Cumming, Burchard, Zeoli, & Ellerby, 2010) and they are more accurate than unstructured risk assessment methods (Bonta, Law, & Hanson, 1998; Dawes et al., 1989; Grove, Zald, Lebow, Snitz, & Nelson, 2000; Hanson & Morton-Bourgon, 2009; Mossman, 1994).

Static-99 (Hanson & Thornton, 2000) is the most commonly used actuarial scale for sex offenders. It is widely used in Canada and the United States for treatment planning (McGrath et al., 2010; Jackson & Hess, 2007), community supervision (Interstate Commission for Adult Offender Supervision, 2007), and civil commitment evaluations (Jackson & Hess, 2007). Although the predictive accuracy of Static-99 is not notably better than other actuarial tools designed for sex offenders (Hanson & Morton-Bourgon, 2009), its popularity is probably due to it being easily scored by diverse professionals using widely available information (e.g., criminal history, offender age). The developers of Static-99 also created a new scale, Static-2002, which was intended to have increased coherence and conceptual clarity (see Hanson & Thornton, 2003). In a multisite study across eight diverse samples, Static-2002 had significantly greater accuracy than Static-99 in predicting sexual, violent, and any recidivism, although the difference for sexual recidivism was quite small (Hanson, Helmus, & Thornton, 2010).

Though existing actuarial risk scales are superior to unstructured clinical risk assessment, they usually cannot take into account all potentially relevant risk factors. For example, several studies have found that including a range of external risk factors adds incremental predictive validity to Static-99 (Allan, Grace, Rutherford, & Hudson, 2007; Beech, Friendship, Erikson, & Hanson, 2002; Craig, Thornton, Beech, & Browne, 2007; Hanson, Harris, Scott, & Helmus, 2007; Knight & Thornton, 2007; Olver, Wong, Nicholaichuk, & Gordon, 2007; Thornton, 2002).

Additionally, the items selected for inclusion in an actuarial scale (and the weights assigned to them) reflect their empirical relationship to recidivism for the majority of offenders within the development samples, and may not work for minority subsets of offenders (e.g., internet sex offenders, adolescent offenders released as adults). In particular, studies have found that actuarial risk assessment scales, particularly Static-99, may not adequately account for the advanced age of some offenders (Barbaree, Langton, & Blanchard, 2007; Barbaree, Langton, Blanchard, & Cantor, 2009; Hanson, 2006; Thornton, 2006). Consequently, researchers have proposed post-hoc methods for incorporating age information into actuarial scales, including Bayesian adjustments to recidivism estimates (Wollert, 2006), age-corrected actuarial scales (Barbaree et al., 2007, 2009), and age-stratified recidivism tables (Wollert, Cramer, Waggoner, Skelton, & Vess, 2010).

Considering advanced age takes on a vital importance with current demographic trends. The proportion of seniors has been increasing in the general population (Administration on Aging, 2009; Statistics Canada, 2008; Turcotte & Schellenberg, 2007), as well as in prison populations (Boe, Nafekh, Vuong, Sinclair, & Cousineau, 2003). As actuarial risk scales are increasingly being used with an aging sex offender population, it is necessary to further investigate whether the current age item weights in actuarial scales are optimal.

One of the most strongly supported findings in criminology is that involvement in crime declines with age. This trend is generally invariant across gender, jurisdiction, offence types, race, and time periods (Hirschi & Gottfredson, 1983; Sampson & Laub, 2003), and has been observed for sexual recidivism as well (Barbaree & Blanchard, 2008; Hanson, 2002). Experts debate whether this relationship is caused by maturation effects or an enduring propensity for crime in early-starting delinquents, but the pattern is well established. Knowing why age and recidivism are correlated is desirable, but such knowledge is not needed to conclude that incorporating age can improve risk-assessment measures. For the purposes of gauging risk, the utility of a risk factor hinges simply on its empirical relationship to the outcome of interest (e.g., Meehl, 1956).

Rationale for Current Study

Actuarial scales should be continuously reevaluated and revised to reflect advances in knowledge (Dawes et al., 1989). The accumulation of newer and larger samples provides the opportunity to discover risk factors not adequately captured by current actuarial scales and to revise the scales accordingly. Assessing whether existing actuarial scales appropriately capture risk due to age is important given the robust relationship between age and offending, as well as the increased proportion of older offenders in correctional populations.

The primary purpose of the present study was to examine whether Static-99 and Static-2002 adequately capture the relationship between age at release and recidivism (primarily sexual recidivism, although violent recidivism was also examined). Given our finding of incremental predictive validity for age, new age weights for both risk instruments were developed and tested. This study is part of a larger ongoing project to renorm Static-99. Further research on variables moderating absolute recidivism estimates (e.g., recidivism definition, type of sample) will be available in forthcoming manuscripts.

In addition to offering a technical update to the age weights of two of the most commonly used risk assessment scales, a secondary purpose of the study was to contribute to the literature on the age/crime relationship. We had insufficient data to test possible causal explanations for our findings. The large and diverse samples of the current study, however, allowed a more detailed examination of recidivism rates by age cohorts (focusing on offenders above age 40) than has been examined in previous studies. This article, therefore, identifies basic relationships between advanced age

and recidivism that require explanation by theories on the age/crime relationship in sexual offending.

The current study evaluated two risk properties of actuarial scales: relative and absolute risk (also referred to as discrimination and calibration, respectively, by Gail & Pfeiffer, 2005). Relative risk (e.g., risk categories, percentiles, relative risk ratios) provides information about a particular offender's risk relative to other offenders (i.e., comparing recidivists and nonrecidivists) and can be examined with statistics such as AUCs, Cohen's *d*, and B_1 logistic regression coefficients. Relative risk can inform most decisions involving the allocation of scarce resources (e.g., treatment, supervision) and remains fairly consistent across diverse samples (Hanson et al., 2010). Absolute risk, however, refers to the expected probability of recidivism, typically reported as recidivism estimates from survival analysis or logistic regression (B_0). Decisions involving thresholds, such as civil commitment evaluations, typically require some estimate or approximation of absolute risk (e.g., "more likely than not," "significant probability").

Method

Measures

Static-99 (Hanson & Thornton, 2000). Static-99 is an empirically derived actuarial risk assessment tool designed to predict sexual and violent recidivism in adult male sex offenders (see also www.static99.org). It has 10 items and the total score (ranging from 0 to 12) can be used to place offenders in one of four risk categories: low (0-1), moderate-low (2-3), moderate-high (4-5), and high (6+; Harris, Phenix, Hanson, & Thornton, 2003). Static-99 includes one item for age at release, which is scored dichotomously: one risk point is given to offenders less than 25 years old.

Static-2002 (Hanson & Thornton, 2003). Similar to Static-99, Static-2002 is an empirically derived actuarial risk assessment tool designed to predict sexual and violent recidivism in adult male sex offenders (see also www.static99.org). It has 14 items grouped into five main subscales: age at release, persistence of sexual offending, sexual deviance, relationship to victims, and general criminality. The total score (ranging from 0 to 14) can be used to place offenders in one of five risk categories: low (0-2), low-moderate (3-4), moderate (5-6), moderate-high (7-8), and high (9+; Phenix, Doren, Helmus, Hanson, & Thornton, 2009). Static-2002 also has an item for age at release, but it is coded on a four-point scale, with offenders less than 25 receiving three points, offenders between 25 and 34.9 years old receiving two points, offenders between 35 and 49.9 years old receiving one point, and offenders age 50+ receiving zero points.

Samples

Raw data (typically in the form of SPSS datasets) were obtained by contacting the authors of all known Static-99 and Static-2002 replications with adult male offenders.

To include a dataset, we required information on the offender's age at release and sufficient recidivism information (i.e., fixed follow-up outcomes) to conduct logistic regression analyses. In total, 24 Static-99 samples were obtained (23 with sexual recidivism and 19 with violent recidivism data), seven of which also included Static-2002 scores. There are fewer samples in the current study compared to the larger project renorming Static-99 (Helmus, 2009) because age information was required for the current study.

Although it was not possible to evaluate the quality of the Static ratings in the samples, minimal standards for quality control were applied. Cases were deleted if there were unresolved coding inconsistencies. Additionally, as per the coding rules (Harris et al., 2003; Phenix et al., 2009), cases were deleted if more than one Static-2002 item was missing, any Static-99 item was missing other than Ever Lived with a Lover (Item 2), the offender was less than 18 years old at time of release or less than 16 years old when they committed the index offence, or if the offender was female. All data analyses were conducted independently by the first and fourth author to ensure accuracy.

Tables 1 and 2 provide basic descriptive information for the studies included. For more information on the samples, readers are encouraged to obtain a more detailed report of this project (Helmus, 2009; available from www.static99.org) or to refer to the original studies. The total sample included 8,390 sex offenders with Static-99 scores and 2,609 with Static-2002 scores. Eleven samples were from Canada, six were from the United States, two were from the United Kingdom, and there was one each from Denmark, Austria, Sweden, Germany, and New Zealand. Of the 14 studies that could be classified in terms of their treatment status, seven samples were mostly treated (defined as more than 75% of the offenders), whereas five were mixed in their treatment exposure, and only one sample was mostly untreated (less than 25%). One additional sample (G. T. Harris et al., 2003) consisted of two subsamples, one of which was mostly treated and another that was mixed in terms of treatment exposure. The average age at release was 40 years old (SD = 12), ranging between 18 and 84, with only six offenders in their 80s. Offenders were released between 1957 and 2007, although 81% were released in 1990 or later.

All samples used official criminal records to measure recidivism, but 13 samples used charges as the recidivism criteria and 11 used convictions. Although experts disagree on which definition is preferable for recidivism research, previous analyses did not find consistent differences in recidivism rates based on whether these samples used charges or convictions (Helmus, 2009). Either definition underestimates the true rate of recidivism.

Additionally, a methodological artifact common to sexual recidivism research is that offenders can be charged with violent offences that were actually sexual in motivation (e.g., as the result of a plea bargain). To our knowledge, only six samples attempted to determine sexual motivation (Craissati, Bierer, & South, 2008; Hanson et al., 2007; Hill, Haberman, Klusmann, Berner, & Briken, 2008; Knight & Thornton, 2007; Nicholaichuk, 2001; Wilson, Cortoni, & Vermani, 2007; Wilson, Picheca, &

Table 1. Descriptive Information

Study	*n*	Age *M* (*SD*)	Country	Recidivism criteria	Type of sample	Mostly treated	Release period	Md year release
Allan et al. (2007)	492	42 (12)	New Zealand	Charges	Prison treatment	Yes	1990-2000	1994
Bartosh et al. (2003)	186	38 (12)	United States	Charges	Routine correctional		1996	1996
Bengtson (2008)	311	33 (10)	Denmark	Charges	Forensic psychiatric evaluations		1978-1995	1986
Bigras (2007)	483	43 (12)	Canada	Charges	Routine CSC	Mixed	1995-2004	1999
Boer (2003)	299	41 (12)	Canada	Conviction	Routine CSC		1976-1994	1990
Bonta & Yessine (2005)	133	40 (10)	Canada	Conviction	Preselected high risk	Mixed	1992-2004	1999
Brouillette-Alarie & Proulx (2008)	228	36 (10)	Canada	Conviction	Prison & community treatment		1979-2006	1996
Cortoni & Nunes (2007)	73	42 (12)	Canada	Charges	Prison treatment	Yes	2001-2004	2003
Craissati et al. (2008)	209	38 (12)	United Kingdom	Conviction	Routine community supervision	Mixed	1992-2005	1998
Eher et al. (2008, 2009) (2008)	706	41 (12)	Austria	Conviction	Routine European prison		2000-2005	2003
Epperson (2003)	177	37 (13)	United States	Charges	Routine correctional		1989-1998	1995
Haag (2005)	198	37 (10)	Canada	Conviction	Detained until end of sentence	Mixed	1995	1995
Hanson et al. (2007)	702	42 (13)	Canada	Charges	Routine community supervision		2001-2005	2002
Harkins & Beech (2007)	197	43 (12)	United Kingdom.	Convictions	Prison & community treatment	Yes	1994-1998	1995
G.T. Harris et al. (2003)	284	37 (12)	Canada	Charges	High intensity treatment and psychiatric referrals	Yes & Mixed	1968-1996	1988
Hill et al. (2008)	86	39 (11)	Germany	Conviction	Sexual homicide perpetrators		1971-2002	1989
Johansen (2007)	273	38 (11)	United States	Charges	Prison treatment	Yes	1994-2000	1996

(continued)

Table 1. (continued)

Study	*n*	Age *M* (*SD*)	Country	Recidivism criteria	Type of sample	Mostly treated	Release period	Md year release
Knight & Thornton (2007)	466	36 (11)	United States	Charges	Civil commitment evaluation		1957-1986	1970
Långström (2004)	1,278	41 (12)	Sweden	Conviction	Routine European prison	No	1993-1997	1995
Nicholaichuk (2001)	281	35 (9)	Canada	Conviction	High intensity treatment	Yes	1983-1998	1992
Saum (2007)	169	46 (12)	United States	Charges	Community supervision	Yes	1988-1998	-
Swinburne Romine et al. (2008)	680	38 (12)	United States	Conviction	Community treatment	Mixed	1977-2007	1988
Ternowski (2004)	247	44 (13)	Canada	Charges	Prison treatment	Yes	1994-1998	1996
Wilson et al. (2007a & b)	232	42 (11)	Canada	Charges	Preselected high risk		1994-2007	2002
Total	8,390	40 (12)					1957-2007	1995

Note: CSC = Correctional Service Canada (administers all prison sentences of at least 2 years in Canada).

Table 2. Recidivism Information

Study	Static-99 *M* (*SD*)	Static-2002 *M* (*SD*)	Years Follow-Up *M* (*SD*)	Total *N*	5-year *n*	10-year *n*	Sexual recidivism (%)			Violent recidivism (%)		
							Overall	5 year	10 year	Overall	5 year	10 year
Allan et al. (2007)	2.2 (2.0)	—	5.7 (2.9)	492	298	25	9.6	11.7	20.0	16.5	18.1	36.0
Bartosh et al. (2003)	3.5 (2.4)	—	5.0 (0.2)	186	90	—	11.8	13.3	—	24.2	26.7	—
Bengtson (2008)	3.8 (2.1)	5.1 (2.3)	16.2 (4.2)	311	310	291	33.8	19.7	28.5	51.8	32.3	44.0
Bigras (2007)	2.7 (2.0)	4.1 (2.3)	4.6 (1.9)	483	207	—	6.2	9.2	—	15.3	22.2	—
Boer (2003)	3.3 (2.3)	4.5 (2.5)	13.3 (2.1)	299	299	295	8.7	3.7	7.8	23.4	14.0	21.0
Bonta & Yessine (2005)	5.3 (1.9)	—	5.5 (2.4)	133	81	3	15.8	17.3	0.0	33.8	39.5	33.3
Brouillette-Alarie & Proulx (2008)	3.9 (2.2)	—	9.9 (4.5)	228	199	110	20.2	14.6	20.9	30.7	22.1	31.8
Cortoni & Nunes (2007)	2.6 (1.8)	—	4.6 (0.6)	73	17	—	0.0	0.0	—	8.2	11.8	—
Craissati et al. (2008)	2.4 (2.0)	—	9.1 (2.7)	209	200	66	11.5	7.5	9.1	24.4	16.0	24.2
Eher et al. (2008, 2009)	2.7 (2.0)	—	3.9 (1.1)	706	151	—	4.0	2.0	—	14.7	11.9	—
Epperson (2003)	2.7 (2.2)	—	7.9 (2.5)	177	150	36	14.1	10.7	22.2	—	—	—
Haag (2005)	3.9 (2.0)	6.0 (2.2)	7.0 (0.0)	198	198	—	25.3	19.7	—	—	—	—
Hanson et al. (2007)	2.9 (2.0)	4.0 (2.3)	3.4 (1.0)	702	31	—	8.1	0.0	—	16.4	3.2	—
Harkins & Beech (2007)	2.8 (2.2)	4.3 (2.6)	10.4 (1.1)	197	197	127	14.2	9.6	16.5	20.8	13.7	23.6
G.T. Harris et al. (2003)	3.2 (2.4)	—	7.6 (4.2)	284	197	87	—	—	—	34.2	28.9	43.6
Hill et al. (2008)	4.9 (1.8)	—	12.6 (6.6)	86	73	54	15.1	11.0	18.5	29.1	23.3	37.0
Johansen (2007)	3.0 (2.0)	—	9.1 (1.1)	273	272	62	7.7	5.9	12.9	20.5	15.1	17.7
Knight & Thornton (2007)	4.5 (2.2)	6.5 (2.5)	8.6 (2.6)	466	433	353	26.2	24.7	30.0	36.9	32.5	41.3
Långström (2004)	2.4 (2.0)	—	8.9 (1.4)	1,278	1,278	353	7.5	5.4	7.4	21.4	15.3	22.9
Nicholaichuk (2001)	4.7 (2.1)	—	6.4 (4.0)	281	168	59	18.5	22.6	25.4	—	—	—
Saum (2007)	2.0 (1.5)	—	5.0 (0.0)	169	169	—	33.7	29.6	—	—	—	—
Swinburne Romine et al. (2008)	1.9 (1.8)	—	16.8 (7.8)	680	569	542	13.8	8.4	11.3	—	—	—
Ternowski (2004)	2.1 (1.9)	—	7.5 (1.0)	247	247	—	8.1	6.5	—	15.4	13.4	—
Wilson et al. (2007a & b)	5.5 (2.0)	—	5.2 (3.0)	232	103	16	10.3	11.7	6.3	25.9	32.0	43.8
Overall	3.0 (2.2)	4.8 (2.6)	8.2 (5.0)	8,390	5,937	2,479	12.4	11.1	16.6	23.0	20.1	31.8

Note: Recidivism information is from fixed follow-up periods, not controlling for Static scores.

Prinzo, 2007), and the success of these attempts is difficult to discern. For example, Wilson, Cortoni, et al. (2007) and Wilson, Picheca, et al. (2007) knew the circumstances of some of the recidivism incidents but did not routinely have access to this information (R. Wilson, personal communication, June 16, 2010). Hanson and colleagues (2007), however, conducted a more systematic search by contacting the relevant police departments for violent recidivism incidents to assess possible sexual motivation. Of the remaining 18 studies, three specified that reoffence circumstances were not known and 15 made no mention of determining sexual motivation (presumably it was not assessed).

Table 2 presents average Static-99/2002 scores per sample (for Static-99, $M = 3.0$, $SD = 2.2$; for Static-2002, $M = 4.8$, $SD = 2.6$). Offenders were followed up for an average of 8.2 years ($SD = 5.0$). Table 2 also includes sexual and violent recidivism rates overall (not controlling for follow-up time) and rates from fixed 5- and 10-year follow-up periods. These data do not control for Static scores. The observed sexual recidivism rate for all cases was 12.4%, with a 5-year rate of 11.1% and 10-year rate of 16.6%. The observed violent recidivism rate for all cases was 23.9%, with a 5-year rate of 20.7% and 10-year rate of 32.5%. From the original 24 samples ($N = 8,390$), approximately 70% of cases had at least a 5-year follow-up ($n = 5,937$), and only 30% had 10-year follow-up data ($n = 2,479$). More fluctuation across samples is, therefore, observed at 10 years due to smaller sample sizes. In some cases, the smaller subset of offenders with 10-year data had lower 10-year recidivism rates than the 5-year rates from the larger group of offenders.

Offender type information was available in 16 samples. Noncontact offenders (e.g., exhibitionists, voyeurs) and mixed offenders (those with both adult and child victims) were identified inconsistently and in small numbers and were excluded from offender type analyses. The current literature provides no clear consensus on how to classify offenders as rapists or child molesters and most samples did not specify the definition that was applied. When sufficient victim information was available, however, the following definition was used: offenders with victims less than 14 were considered child molesters, and offenders with adult victims (age 18+) were classified as rapists. For offenders with victims between the ages of 14 and 17 inclusive, they were classified as child molesters if their victims were related and as rapists if the victims were unrelated. Offenders with both adult and child victims were classified based on their predominant victim choice (if possible).

Coding Age at Release

Although the weights vary across the age items in Static-99 and Static-2002, the coding rules remain the same (Harris et al., 2003; Phenix et al., 2009). The items are coded based on the offender's age on the day he was released from the most recent offence that forms a part of the index sex offence. It is possible in some cases that the index sex offence identified for scoring purposes is not the same as the offender's current offence (e.g., an offender serving a sentence for a nonsexual offence being

assessed as a sex offender due to a prior sexual offence). In these cases, the offenders may be substantially older at assessment than when they were released from their index sex offence.

In the datasets used for the current study, only one sample specified if the offender's current offence was different from the index sex offence (Bonta & Yessine, 2005). In this sample, cases were deleted if the index sex offence was more than 2 years prior to the current offence ($n = 22$) because Static-99 and Static-2002 were developed on, and intended for, sexual offenders with a current (or recent) sexual offence. The remaining studies did not identify cases where the index sex offence was not the same as the current offence; it is presumed that the number of such cases would be small. For all cases, the offender's date of birth and release date were used to check that the age at release item was correctly coded.

Overview of Analyses

Actuarial risk prediction tools provide numeric estimates of both relative and absolute risk. To evaluate differences between recidivists and nonrecidivists (discrimination, or relative risk), we used areas under the receiver operating characteristic curves (ROC AUC), slope coefficients from logistic regression, and hazard rates from Cox regression. Each of these statistics has its own strengths and weaknesses. The AUC is useful for comparing results across samples because it is not influenced by the outcome variable (i.e., recidivism base rates; Humphreys & Swets, 1991; Rice & Harris, 2005). It is, however, influenced by the variance in the scores used to predict recidivism (Hanson, 2008; Humphreys & Swets, 1991). In other words, even when the observed recidivism rates per score are identical across studies, AUCs may change depending on the distribution of risk assessment scores. All else being equal, AUCs will be smaller in samples preselected on risk (all high risk or all low risk) than when the AUCs are computed using the complete sample (a mixture of high-risk and low-risk offenders). We used the Hanley and McNeil (1983) test of correlated AUC areas to test for differences in relative predictive accuracy. This method has been used in other studies comparing the predictive accuracy of the RRASOR, Static-99/R, and Static-2002/R (Babchishin, Hanson, & Helmus, 2011; Hanson et al., 2010; Hanson & Thornton, 2000).

Logistic regression and Cox regression were also used because they tend to provide more stable estimates of relative predictive accuracy (i.e., are less influenced by restriction of range; Hanson, 2008). Cox regression estimates relative risk ratios (hazard rates) associated with one or more predictor variables from survival data with unequal follow-up times (Allison, 1984). One disadvantage of Cox regression is that it assumes the rates of recidivism over time (i.e., the shape of the survival curve) are approximately the same across all samples and all risk levels. This assumption is unsupported based on our findings of considerable variability in recidivism rates across samples, even after controlling for risk (Helmus, 2009). In our Cox regression analyses, we used sample as a strata variable, which relaxes the assumption, thereby

controlling for sample differences in base rates. This method would also control for other differences across samples, such as recidivism definition (e.g., charges vs. convictions). Cox regression does not provide estimates of absolute risk.

Logistic regression (Hosmer & Lemeshow, 2000) can be used to examine both relative and absolute risk. Logistic regression is a form of regression in which the dichotomous dependent variable (recidivism) is transformed into log odds. With one predictor variable (Static-99 or Static-2002), logistic regression estimates two regression coefficients (B_0 and B_1). B_1 is an estimate of predictive accuracy, or the average change in recidivism rates for each one-unit increase in risk scores (expressed as a log odds ratio); and B_0 is an estimate of the recidivism base rate for offenders with a score of "zero." To link the logistic regression recidivism estimates to any specific time period, fixed follow-up periods are required (e.g., for a fixed 5-year recidivism outcome, only offenders with at least 5 years of follow-up were included, with recidivism occurring after 5 years excluded; this creates an equal length of follow-up for each offender). Consequently, the sample size available for logistic regression is smaller than for statistics that allow varying follow-up (e.g., Cox regression).

To control for base rate differences across samples in the logistic regression analyses, we used a dichotomous covariate that identified samples as either "routine" or "other," which explained the most variability across samples in recidivism rates (Helmus, 2009).[1] Logistic regression also assumes that the relationship between the predictors and the outcome follows a logistic distribution. Preliminary analyses suggested this assumption was generally met.

For examining absolute recidivism estimates, life table survival analysis (Soothill & Gibbens, 1978) was used. This is the same method used to produce the original Static-99 recidivism estimates (Hanson & Thornton, 2000). Survival analysis corrects for unequal follow-up times by examining recidivism patterns over time and estimating the expected recidivism rate of the sample for any specified time period. In other words, for a specific follow-up period, it estimates what the recidivism rate for the sample would be if everyone had been followed up to that point. Estimates from survival analysis are generally reliable provided that there are a sufficient number of cases with follow-up data for the time period being estimated. Because the estimates are generated independently for each combination of score and follow-up period, estimates tend to be unstable in some cells with small sample sizes (e.g., high-risk offenders above the age of 60). For the purposes of the current study, however, survival analysis provides a simple and useful way to display recidivism rates, although we do not recommend these rates be used in applied reports. For applied reports, we believe that recidivism estimates generated from logistic regression analyses are the most reliable (Hanson et al., 2010). For the Static scales, estimates generated from meta-analysis of logistic regression coefficients are available for different sample types (see www.static99.org).

To evaluate the correspondence between the predicted and the observed recidivism rates (i.e., calibration, or absolute risk) we used the *E*/*O* index (Gail & Pfeiffer, 2005; Rockhill, Byrne, Rosner, Louie, & Colditz, 2003). In this study, the *E*/*O* index was

defined as the ratio of the predicted number of recidivists (E) divided by the observed number of recidivists (O; Method M_0 from Viallon, Ragusa, Chavel-Chapelon, & Bénichou, 2009). Although more complicated estimators of the E/O index are available, this simple E/O estimator is unbiased when complete follow-up information is available (recidivism was examined for 5-year fixed follow-up periods, with predicted recidivism rates calculated using logistic regression). Following Rockhill et al. (2003), the 95% confidence intervals for the E/O index were calculated using the Poisson variance for the logarithm of the observed number of cases (O):

$$95\%\,\mathrm{CI}(E/O) = (E/O)\exp\left(\pm 1.96\sqrt{1/O}\right)$$

As an overall significance test for sets of predicted values, we used the traditional chi-square goodness of fit statistic (Ferguson, 1976; Equation 13.1):

$$\chi^2 = \sum \frac{(O-E)^2}{E}$$

which is tested against a chi-square distribution with degrees of freedom equal to one less than the number of expected values compared. The E/O index can be considered a measure of effect size for the fit within a group, whereas the chi-square is a test of significance and can be calculated for the overall fit between observed and expected values across multiple groups. Perfect fit is indicated by an E/O index of 1.0, and confidence intervals that do not include 1.0 indicate significant differences between the observed and predicted recidivism rates.

For the E/O statistics, the observed number of recidivists at 5 years was compared to the predicted number of recidivists from logistic regression analyses at 5 years. In logistic regression, expected probabilities are calculated in the same way as for traditional regression analyses, where the predicted value for a given score on the independent variable (Static-99) is obtained by adding the intercept (the predicted value for a score of 0; the B_0) to the product of the slope and the independent variable. The only difference is that the predicted values are in the unit of logits (i.e., the natural logarithm of the odds, $p/(1-p)$]. Logits were then transformed back into probabilities, where

$$\text{Probability} = \frac{e^{LOGIT}}{1+e^{LOGIT}}$$

Results

Age was negatively correlated with Static-99 ($r = -.222$, $p < .001$) and Static-2002 scores ($r = -.405$, $p < .001$). When the age item was removed from the scales, the correlation decreased in magnitude but remained negative and significant (for Static-99, $r = -.157$, $p < .001$; for Static-2002, $r = -.087$, $p < .001$). Negative

Table 3. Estimated 5-Year Recidivism Rates (Survival Analysis) by Static-99 Risk Category and Age Band

	18-30		30-39.9		40-49.9		50-59.9		60-69.9		70+	
Static-99 risk category	Initial *n*	Recid (%)	Initial *n*	Recid (%)	Initial *n*	Recid (%)	Initial *n*	Recid (%)	Initial *n*	Recid (%)	Initial *n*	Recid (%)
Sexual recidivism												
Low	185	5.8	733	5.5	759	3.6	417	3.3	218	1.9	68	1.5
Moderate-low	672	8.5	855	7.5	666	5.6	315	5.9	140	3.1	37	6.7
Moderate-high	674	15.1	568	18.4	382	9.8	140	12.9	50	7.9	16	0.0
High	328	28.6	390	28.8	297	23.6	127	20.1	51	10.9	18	5.6
Total	1,859	14.2	2,546	12.5	2,104	8.0	999	7.4	459	3.8	139	3.2
Violent recidivism												
Low	145	14.6	594	11.3	620	6.2	367	4.5	178	0.0	59	2.9
Moderate-low	548	21.6	731	16.9	559	12.9	278	7.7	116	4.8	33	4.4
Moderate-high	570	35.9	485	36.1	336	17.1	135	18.9	48	10.2	14	10.5
High	284	49.0	330	42.8	271	31.0	120	27.3	48	17.6	16	12.5
Total	1,547	31.4	2,140	23.6	1,786	14.0	900	10.5	390	4.8	122	5.5

Note: Recid = recidivism. These recidivism estimates were obtained from survival analysis and are presented for illustrative purposes only. For applied risk assessment reports, readers are encouraged to obtain the most recent recidivism estimates generated from logistic regression analyses (available from www.static99.org).

Table 4. Estimated 5-Year Sexual Recidivism Rates (Survival Analysis) by Static-2002 Risk Category and Age Band

	18-30		30-39.9		40-49.9		50-59.9		60+	
Static-2002 risk category	Initial *n*	Recid (%)	Initial *n*	Recid (%)	Initial *n*	Recid (%)	Initial *n*	Recid (%)	Initial *n*	Recid (%)
Sexual recidivism										
Low	10	0.0	118	4.8	59	4.7	136	2.3	108	1.9
Low-moderate	130	8.2	245	6.9	204	5.7	74	4.6	40	2.7
Moderate	267	14.9	232	13.2	142	11.9	51	12.2	22	4.6
Moderate-high	157	23.6	150	27.3	91	27.1	29	23.1	17	13.1
High	96	31.7	78	37.4	40	25.4	6	16.7	7	14.3
Total	660	18.1	823	15.0	636	11.4	296	7.2	194	4.0
Violent recidivism										
Low	10	0.0	114	7.4	156	7.0	134	6.6	105	2.8
Low-moderate	125	15.2	233	15.6	194	15.4	71	9.0	40	5.2
Moderate	244	37.7	205	28.7	129	22.4	45	18.6	21	4.9
Moderate-high	139	40.1	123	43.9	81	32.0	27	27.2	16	14.1
High	90	49.2	68	43.5	36	31.0	6	16.7	7	14.3
Total	608	35.2	743	25.3	596	17.9	283	11.5	189	5.0

Note: Recid = recidivism. These recidivism estimates were obtained from survival analysis and are presented for illustrative purposes only. For applied risk assessment reports, readers are encouraged to obtain the most recent recidivism estimates generated from logistic regression analyses (available from www.static99.org).

correlation coefficients indicate that older offenders score lower in static risk, even though they have had more time to accumulate a criminal history than younger offenders.

Are Age Adjustments Needed?

Tables 3 and 4 present the estimated sexual and violent recidivism rates (generated from survival analysis) by age group (<30, 30-39.9, 40-49.9, 50-59.9, 60-60.9, and 70+) and Static risk category. For Static-99 (Table 3), sexual and violent recidivism rates predictably increased with each Static-99 risk category. The only exception to this pattern is among the 70+ group, where the overall sexual recidivism rate was generally low (3.2%) and there was substantial fluctuation across risk categories due to small sample sizes. Generally, within each risk category, sexual and violent recidivism rates decreased for each age group, with minor fluctuations observed. For Static-2002 (Table 4), the 70+ group was combined with the 60-69.9 group because of small sample sizes. Similar patterns emerged, with minor fluctuations in some cells with small sample sizes. The results from both Static-99 and Static-2002 demonstrate a clear decline in recidivism among older sexual offenders, even within the same risk category (although the age declines were less marked for Static-2002).

Cox regression was used to test the incremental predictive validity of age at release for Static-99 and Static-2002. For Static-99, analyses of sexual recidivism included 8,025 cases from 22 samples (Cortoni & Nunes, 2007, was excluded because using sample as strata requires at least one recidivist per sample). After controlling for Static-99, age at release had a significant negative relationship with sexual recidivism (χ^2 change = 31.2, $df = 1$, $p < .001$). The hazard ratio was 0.98 (95% CI of 0.98 to 0.99), meaning that each 1-year increase in age was associated with 98% of the recidivism rate from the previous (younger) age. In other words, the expected recidivism rate of 32-year-old offenders is 98% of the recidivism rate of 31-year-old offenders, which is 98% of the recidivism rate of 30-year-old offenders, and so on. Nonlinearity was tested by entering a squared (one curve) age at release variable in addition to the original (linear) age variable (and Static-99 score). The results were significant, (χ^2 change = 10.5, $df = 1$, $p = .001$), indicating a nonlinear relationship between age and recidivism. Adding a cubed (two curve) age at release variable was nonsignificant (χ^2 change = 2.7, $df = 1$, $p = .098$).

Similar results were found for violent recidivism, although the effect of age was stronger. After controlling for Static-99, the hazard ratio for age was 0.97 (95% CI of 0.96 to 0.97), indicating that the recidivism rate of each age group was 97% of the rate of offenders one year younger (χ^2 change = 217.7, $df = 1$, $p < .001$). The relationship was nonlinear, with a significant incremental quadratic (χ^2 change = 11.0, $df = 1$, $p = .001$) and cubic effect (χ^2 change = 4.1, $df = 1$, $p = .043$). This indicated that the curvilinear relationship between age and violent recidivism was best described with two curves.

After controlling for Static-2002 scores ($k = 7$, $n = 2{,}606$), the linear effect of age for the prediction of sexual recidivism was nonsignificant, (χ^2 change = 2.0, $df = 1$, $p = .162$),

but there was a significant quadratic effect (χ^2 change = 6.7, df = 1, p = .010). The cubic effect was nonsignificant (χ^2 change = 0.01, df = 1, p = .926). For violent recidivism, age had a stronger effect after controlling for Static-2002 scores (hazard ratio of .98, 95% CI of 0.97 to 0.98; χ^2 change = 42.9, df = 1, p < .001), with a significant quadratic effect (χ^2 change = 4.9, df = 1, p = .027) and nonsignificant cubic effect (χ^2 change = 0.6, df = 1, p = .431). For both recidivism outcomes, the relationship between age and recidivism (after controlling for Static-2002) was best represented by a line with one curve.

The incremental effects of age indicated that neither Static-99 nor Static-2002 sufficiently accounted for age at release for sexual or violent recidivism. Adjustment of the scales was, therefore, necessary. The observed nonlinear effects, however, may be misleading given that age items were already included in the scales. Rerunning the Cox regression analyses for sexual recidivism while controlling for Static-99 and Static-2002 scores computed without the age item, the linear effect of age was still significant (for Static-99, χ^2 change = 52.5, df = 1, p < .001; for Static-2002, χ^2 change = 26.3, df = 1, p < .001), but the quadratic effect was not (for Static-99, χ^2 change = 2.4, df = 1, p = .124; for Static-2002, χ^2 change = 3.0, df = 1, p = .082). For violent recidivism, similar results were found, with a linear effect of age (for Static-99, χ^2 change = 278.8, df = 1, p < .001; for Static-2002, χ^2 change = 117.9, df = 1, p < .001), but not a quadratic effect (for Static-99, χ^2 change = 2.0, df = 1, p = .157; for Static-2002, χ^2 change = 1.8, df = 1, p = .178).

Interactions between age and Static scores were also tested to determine whether the effect of age differed across risk levels (e.g., it may be possible that high-risk offenders persist in offending regardless of age). For sexual recidivism, controlling for age at release and Static scores (computed without the age items), the interactions were nonsignificant (for Static-99, χ^2 change = 1.3, df = 1, p = .248; for Static-2002, χ^2 change = 1.9, df = 1, p = .169). These findings provided no empirical basis for supposing the relationship between age and sexual recidivism is different for persistent high-risk offenders. For violent recidivism, however, the interactions reached significance for Static-99 (χ^2 change = 6.2, df = 1, p = .013) and approached significance for Static-2002 (χ^2 change = 3.0, df = 1, p = .085). Inspection of the data revealed age declines in recidivism for all risk categories, but there was a slightly larger age effect for lower risk offenders. Given that the Static scales were designed specifically to predict sexual recidivism, the revisions were based on the patterns observed for sexual recidivism (i.e., no interaction between age and risk). The analyses suggested that age adjustments could proceed on the assumptions that the relationship between age and violent or sexual recidivism was linear, and there was no notable interaction between age and risk that needed to be accounted for.

Developing and Testing a New Age Item

We believed that revising the age weights for both scales was the simplest way to incorporate the age effect, rather than developing post-hoc age adjustments. We used Static-99 data to create the new age weights for both scales because the sample size

was larger and we assumed the weights would apply equally to Static-99 and Static-2002. The overall Static-99 sample with sexual recidivism data (k = 23, N = 8,106) was divided into two subsamples. Offenders with follow-up periods of less than 10 years (k = 23, n = 5,714) were used as the development sample, and offenders with follow-up periods greater than 10 years were retained for validation (k = 15, n = 2,392). This method allowed the new age weights (developed using 5-year recidivism information) to be validated with another sample at 5 years, and also permitted testing of its generalization to 10-year recidivism rates. Although validating the weights on data with at least a 10-year follow-up would disproportionately select offenders released earlier, this is unlikely to affect the results because previous analyses did not find consistent cohort effects (Helmus, 2009).

The development dataset was analyzed independently by the second and third authors, who conferred before producing a final recommendation. Neither analyst had access to the validation sample until after a consensus recommendation had been announced to the other team members.

The principles guiding the selection of the new age weights were as follows: (a) the units should be integers (whole numbers); (b) each unit should approximate the Static-99 units found in other analyses (odds ratio, risk ratio ≈ 1.33); (c) offenders with the median age (39 years old) should receive a score of 0 for the age item; (d) the overall trend should be a relatively gradual decline in risk (i.e., abrupt changes between adjacent categories were interpreted as random error); (e) the revised Static-99 and Static-2002 measures should have higher overall predictive accuracy than the original (specifically, absolute accuracy should be higher; relative accuracy may not change substantially but should not be worsened); and (f) age should no longer contribute significantly once the new age items of Static-99 and Static-2002 are included. Although stronger age effects were found for violent recidivism, the new age weights were developed based on the sexual recidivism data (our principal outcome of interest).

Several different analytic techniques informed the new age weights. The simplest approach involved examining the 5-year sexual recidivism rates (fixed follow-up) after grouping the offenders into 5-year age bands (with the exception of a 7-year age band for those 18 to 24.9 years old, and a single category for offenders 75 years old or older). The observed recidivism rate for each age band was then compared to the observed 5-year sexual recidivism for the 35 to 39.9 age band (10.3%). The difference was expressed in Static-99 relative risk units (i.e., 1 unit = 1.33 change in relative risk). For example, a recidivism rate of 13.7% would equal a one unit increase (10.3% × 1.33), and a recidivism rate of 5.82% would equal a two unit decrease (10.3% × 1.33^{-2}). The weights suggested by the univariate relationship with age were larger (more extreme) than those finally recommended because they ignored the negative correlation between age and risk scores.

Logistic regression and Cox regression were also used to estimate odds ratios and risk ratios, respectively, after controlling for the total Static-99 score (without the age item). The relative risk indicators (e^B) were then translated into the closest integer

Table 5. Age Weights in Static-99, Static-99R, Static-2002, and Static-2002R

	Measure			
Age band	Static-99	Static-99R	Static-2002	Static-2002R
18-24.9	1	1	3	2
25-34.9	0	1	2	2
35-39.9	0	0	1	1
40-49.9	0	−1	1	0
50-59.9	0	−1	0	0
60+	0	−3	0	−2
Full range	0, +1	−3 to +1	0 to +3	−2 to +2

value of Static-99 relative risk units. For example, an age band associated with a hazard (rate) ratio of 1.13 was given a weight of zero (relative risk rounding to 1), whereas an age band associated with a rate ratio of 1.24 was given a weight of 1 (relative risk rounding to 1.33).

Each of the different statistical approaches to estimating age weights suggested similar, although not identical, results. Not surprisingly, the age weights suggested by the independent authors were similar, but not identical. Specifically, the weights suggested were as follows: (a) 18-34.9 = 2, 35 to 59.9 = 0, 60+ = –3; versus (b) 18-34.9 = 1; 35 to 39.9 = 0; 40 to 59.9 = –1, 60+ = –3. Of these two proposals, the latter was better on 6 of 7 statistical indicators examining Static-99 (with the new age weight) predicting sexual recidivism in the development sample. All differences, however, were trivial (e.g., Cox regression risk ratios of 1.342 vs. 1.326).

Based on the above analyses, the consensus recommendation was to use the latter coding: specifically, offenders less than 35 would receive 1 point, offenders age 35 to 39.9 would receive 0 points, offenders age 40 to 59.9 would have 1 point subtracted, and offenders age 60 and older would have 3 points subtracted. Although Static-2002 was not used in the development of the new age weights, we applied the same age weights to both Static-99 and Static-2002 as there was no justification that the weights should be different. For Static-2002, however, a constant of 1 point was added to maintain consistency with the previously established risk categories. The revised Static scales with these new age weights were called Static-99R and Static-2002R. For Static-99R, the total scores range from –3 to 12 (instead of 0-12 in the original scale; Static-99R $M = 2.7$, $SD = 2.6$), and for Static-2002R, the scores range from –2 to 13 (instead of 0-14 in the original scale; Static-2002R $M = 4.3$, $SD = 2.7$). Table 5 summarizes the age weights in the original and revised scales.

Relative Risk

Due to the small proportion of offenders above 50 and because the new age weights reduce variability in scores for older offenders (they score lower on the revised

scales), we expected the revisions to have a smaller effect on relative predictive accuracy (discrimination) compared to absolute predictive accuracy (calibration). In other words, only small changes in the overall AUCs were expected; we did expect, however, that the absolute recidivism rates from the revised scales would be a significantly better fit for older offenders than the predicted values of the original scales.

Comparing Static-99R to Static-99 in the validation sample, there was a slight increase in relative predictive accuracy for Static-99R for sexual recidivism, as measured by the AUC values using fixed follow-up periods (at 5 years, AUC for Static-99R was .720, compared to .713 for Static-99; at 10 years, AUC for Static-99R was .710, compared to .706 for Static-99). The differences between the original and revised scales for sexual recidivism were not statistically significant. For violent recidivism, greater increases in relative predictive accuracy were observed, with Static-99R demonstrating significantly greater predictive accuracy than Static-99 (at 5 years, AUC for Static-99R was .715, compared to .695 for Static-99, difference = .020, 95% CI of .002 to .038; at 10 years, AUC for Static-99R was .713, compared to .692 for Static-99, difference = .021, 95% CI of .005 to .037). Additionally, no statistical shrinkage (whereby effect sizes tend to decrease in independent replications) was observed with Static-99R. The AUC for 5-year sexual recidivism was .720 in the validation sample compared to .708 in the construction sample.

Table 6 displays the results of logistic regression (at 5 and 10 years) and Cox regression analyses in the validation sample. In all three analyses, age at release did not add significant predictive accuracy after controlling for Static-99R, whereas it did add incrementally to the original Static-99. For violent recidivism, however, age continued to add incrementally in all analyses of both Static-99 and Static-99R, although the effect of age was reduced in the revised scale compared to the original scale (see Table 6). In these analyses, none of the quadratic effects for age were statistically significant.

Static-2002R and Static-2002 were compared in the seven samples with Static-2002 information. All data were examined together (as opposed to the validation sample only) because of the smaller sample size and the absence of shrinkage observed for the new age weights with Static-99R. There was a slight increase in the AUCs for Static-2002R for sexual recidivism (at 5 years, AUC for Static-2002R was .713, compared to .709 for Static-2002; at 10 years, AUC for Static-2002R was .690, compared to .689 for Static-2002). For violent recidivism, there was virtually no difference between the original and revised scales (at 5 years, AUC for Static-2002R was .702, compared to .700 for Static-2002; at 10 years, AUC for Static-2002R was .699, compared to .700 for Static-2002). None of the differences in predictive accuracy between scales were statistically significant.

In the logistic and Cox regression analyses (Table 7), both linear and curvilinear age effects were examined because the preliminary analyses found a curvilinear effect only (not a linear effect). In all analyses, age at release did not add significant predictive accuracy for sexual recidivism after controlling for Static-2002R, whereas the curvilinear effect of age did add incrementally to the original Static-2002 when

Table 6. Comparing Static-99R to Static-99 in the Validation Sample

	Static-99R					Static-99				
	χ^2 change	*df*	*p*	Exp(*B*)	95% CI	χ^2 change	*df*	*p*	Exp (*B*)	95% CI
Sexual recidivism										
Logistic regression 5 years										
Static scale	136.28	1	<.001	1.34	1.27-1.41	128.66	1	<.001	1.36	1.29-1.44
Age	1.13	1	.288	1.01	0.99-1.02	4.19	1	.041	0.99	0.98-0.99
Logistic regression 10 years										
Static scale	164.79	1	<.001	1.33	1.27-1.39	157.14	1	<.001	1.35	1.29-1.42
Age	1.32	1	.251	1.01	0.99-1.02	5.58	1	.018	0.99	0.98-0.99
Cox regression										
Static scale	134.14	1	<.001	1.26	1.21-1.31	123.81	1	<.001	1.28	1.22-1.33
Age	2.15	1	.142	1.01	0.99-1.02	3.44	1	.064	0.99	0.98-1.00
Violent recidivism										
Logistic regression 5 years										
Static scale	149.37	1	<.001	1.33	1.27-1.40	113.88	1	<.001	1.31	1.25-1.38
Age	17.82	1	<.001	0.98	0.96-0.99	54.83	1	<.001	0.96	0.95-0.97
Logistic regression 10 years										
Static scale	187.34	1	<.001	1.33	1.28-1.40	141.10	1	<.001	1.32	1.26-1.38
Age	24.38	1	<.001	0.97	0.96- 0.98	74.80	1	<.001	0.96	0.95-0.97
Cox regression										
Static scale	182.62	1	<.001	1.25	1.21-1.29	135.56	1	<.001	1.23	1.19-1.28
Age	22.85	1	<.001	0.98	0.97-0.99	67.74	1	<.001	0.97	0.96-0.98

Note: CI = confidence interval. Values represent incremental contribution after controlling for previously entered variables. The logistic regression analyses began by controlling for routine/non-routine samples (not shown), then Static, and then age (for sexual recidivism, *n* = 2,392; for violent recidivism, *n* = 1,849). For the Cox regression analyses, samples were considered strata (for sexual recidivism, *n* = 2,388; for violent recidivism, *n* = 1,849).

Table 7. Comparing Static-2002R to Static-2002

	Static-2002R					Static-2002				
	χ^2 change	*df*	*p*	Exp (*B*)	95% CI	χ^2 change	*df*	*p*	Exp (*B*)	95% CI
Sexual recidivism										
Logistic regression 5 years										
Static scale	90.02	1	<.001	1.30	1.23-1.38	84.98	1	<.001	1.30	1.23-1.38
Age	1.83	1	.176	0.99	0.98-1.00	3.40	1	.065	0.99	0.97-1.00
Age^2	0.52	1	.470	1.00	0.99 -1.00	2.85	1	.092	0.99	0.99-1.00
Logistic regression 10 years										
Static scale	65.55	1	<.001	1.27	1.19-1.35	61.96	1	<.001	1.27	1.19-1.35
Age	0.04	1	.847	0.99	0.98-1.01	0.37	1	.542	0.99	0.98-1.01
Age^2	1.65	1	.199	0.99	0.99-1.00	4.54	1	.033	0.99	0.99-0.99
Cox regression										
Static scale	134.69	1	<.001	1.26	1.21-1.31	127.62	1	<.001	1.26	1.21-1.31
Age	0.62	1	.431	0.99	0.99-1.01	1.95	1	.162	0.99	0.98-1.00
Age^2	1.72	1	.190	0.99	0.99-1.00	6.69	1	.010	0.99	0.99-0.99
Violent recidivism										
Logistic regression 5 years										
Static scale	113.21	1	<.001	1.29	1.23-1.36	108.77	1	<.001	1.30	1.23-1.37
Age	30.22	1	<.001	0.97	0.95-0.98	35.80	1	<.001	0.96	0.95-0.98
Age^2	0.01	1	.918	1.00	0.99-1.00	0.75	1	.386	1.00	0.99-1.00
Logistic regression 10 years										
Static scale	103.00	1	<.001	1.30	1.23-1.37	101.05	1	<.001	1.31	1.24-1.38
Age	28.86	1	<.001	0.96	0.95-0.98	33.50	1	<.001	0.96	0.95-0.98
Age^2	0.46	1	.497	1.00	0.99-1.00	0.05	1	.821	1.00	0.99-1.00
Cox regression										
Static scale	202.27	1	<.001	1.25	1.22-1.29	193.34	1	<.001	1.26	1.22-1.30
Age	35.08	1	<.001	0.98	0.97-0.98	42.86	1	<.001	0.98	0.97-0.98
Age^2	0.64	1	.422	1.00	0.99-1.00	4.92	1	.033	0.99	0.99-0.99

Note: CI = confidence interval. Values represent incremental contribution after controlling for previously entered variables. The logistic regression analyses began by controlling for routine/non-routine samples (not shown), then Static, and then age (for sexual recidivism, *n* = 1,644 at 5 years, *n* = 1,057 at 10 years; for violent recidivism, *n* = 1,461 at 5 years, *n* = 1,067 at 10 years). For the Cox regression analyses, samples were considered strata (for sexual recidivism, *n* = 2,606; for violent recidivism, *n* = 2,417).

examining the 10-year logistic regression analysis and the Cox regression analysis. For violent recidivism, however, linear age continued to add incrementally in all analyses of both Static-2002 and Static-2002R, although the effect of age was reduced in the revised scale compared to the original (see Table 7).

The new age weights were also tested to ensure they were appropriate for both rapists and child molesters. In both Cox and logistic regression analyses of sexual and violent recidivism, the interaction between age and offender type was not statistically significant (i.e., predictive accuracy does not differ by offender type; table available on request). There were insufficient data to meaningfully examine the age weights by offender type with Static-2002R.

Absolute Risk

Tables 8 and 9 present estimated sexual recidivism rates (from survival analysis) separated by age band and by Static-99R and Static-2002R risk category. (For both Static-99R and Static-2002R, the same cut-off scores for the nominal risk categories were retained. The proportion of offenders in each risk category and the associated recidivism rates, for both the original and revised scales are presented in appendix).

For Static-99R, sexual recidivism rates per risk category were fairly similar for offenders in their 20s, 30s, 40s and 50s, with minor fluctuations. Approximately 85% of offenders in their 60s and 70s were in the low-risk category, with too few cases in the other categories for meaningful observations (although moderate-low- and moderate-high-risk offenders in their 60s appeared to have higher recidivism rates than similar-aged low-risk offenders). Low-risk offenders in their 60s and 70s appeared to have slightly lower recidivism rates than their younger counterparts. This difference, however, is likely due to variation in average Static-99R scores even within the low-risk category, ($M = -1.6$ for the 70s group, –1.5 for the 60s group, 0.0 in the 50s group, and 0.2 for offenders less than 50). Table 8 also presents separate recidivism estimates for two subgroups within the low-risk category (scores less than zero compared to scores of zero and one).

For violent recidivism rates, consistent with the earlier finding that age still added incrementally to Static-99R, there was a decline in recidivism rates for older age groups within most risk categories (with some fluctuations), although the pattern was less marked than for the original Static-99.

For Static-2002R (Table 9), offenders age 70 and older were combined with offenders in their 60s due to small sample sizes. Also note that for all groups above 50, sample sizes were small in all but the low-risk category. Offenders in their 50s and older had slightly lower sexual recidivism rates than younger offenders. Age declines per risk category were somewhat larger for violent recidivism, although not as marked as for the original Static-2002.

The sexual recidivism rates predicted by Static-99 per age group were significantly different from the observed rates ($\chi^2 = 27.37$, $df = 5$, $p < .001$), whereas Static-99R predicted recidivism rates were not ($\chi^2 = 3.04$, $df = 5$, $p = .694$). Table 10 presents the

Table 8. Estimated 5-Year Recidivism Rates (Survival Analysis) by Static-99R Risk Category and Age Band

	18-30		30-39.9		40-49.9		50-59.9		60-69.9		70+	
Static-99R risk category	Initial *n*	Recid (%)	Initial *n*	Recid (%)	Initial *n*	Recid (%)	Initial *n*	Recid (%)	Initial *n*	Recid (%)	Initial *n*	Recid (%)
Sexual recidivism												
Low	68	5.0	529	5.6	1,130	3.6	603	4.4	391	2.4	115	2.9
Scores < 0	—	—	—	—	362	2.0	205	4.1	306	2.4	91	2.6
Scores of 0 or 1	68	5.0	529	5.6	768	4.4	398	4.6	85	2.7	24	4.3
Moderate-low	575	7.7	844	6.9	528	8.0	214	7.8	34	14.0	14	8.0
Moderate-high	771	12.3	657	14.0	263	17.7	110	12.4	27	8.2	9	0.0
High	445	27.3	516	27.4	183	23.4	72	25.6	7	15.4	1	0.0
Total	1,859	14.2	2,546	12.5	2,104	8.0	999	7.4	459	3.8	139	3.2
Violent recidivism												
Low	54	15.6	435	10.4	930	6.9	530	5.4	325	2.4	100	3.0
Scores < 0	—	—	—	—	291	5.2	179	5.0	252	1.3	79	4.0
Scores of 0 or 1	54	15.6	435	10.4	639	7.6	351	5.6	73	6.2	21	0.0
Moderate-low	465	20.2	712	15.0	452	16.7	197	11.4	32	14.6	14	20.3
Moderate-high	644	31.0	559	30.7	239	29.1	104	21.8	26	20.0	7	14.3
High	384	47.2	434	42.4	165	27.1	69	32.7	7	15.4	1	0.0
Total	1,547	31.4	2,140	23.6	1,786	14.0	900	10.5	390	4.8	122	5.5

Note: Recid = recidivism. These recidivism estimates were obtained from survival analysis and are presented for illustrative purposes only. For applied risk assessment reports, readers are encouraged to obtain the most recent recidivism estimates generated from logistic regression analyses (available from www.static99.org).

Table 9. Estimated 5-Year Sexual Recidivism Rates (Survival Analysis) by Static-2002R Risk Category and Age Band

	18-30		30-39.9		40-49.9		50-59.9		60+	
Static-2002R risk category	Initial *n*	Recid (%)	Initial *n*	Recid (%)	Initial *n*	Recid (%)	Initial *n*	Recid (%)	Initial *n*	Recid (%)
Sexual recidivism										
Low	22	4.6	118	4.8	262	4.6	136	2.3	148	2.1
Low-moderate	197	12.0	245	6.9	180	7.9	74	4.6	22	4.6
Moderate	236	14.5	232	13.2	114	18.4	51	12.2	17	13.1
Moderate-high	125	25.4	150	27.3	59	29.8	29	23.1	6	16.7
High	80	34.0	78	37.4	21	28.6	6	16.7	1	0.0
Total	660	18.1	823	15.0	636	11.4	296	7.2	194	4.0
Violent recidivism										
Low	22	4.6	114	7.4	256	10.8	134	6.6	145	3.4
Low-moderate	188	24.0	233	15.6	167	14.9	71	9.0	21	4.9
Moderate	213	37.6	205	28.7	99	27.9	45	18.6	16	14.1
Moderate-high	110	42.0	123	43.9	55	36.3	27	27.2	6	16.7
High	75	51.5	68	43.5	19	31.6	6	16.7	1	0.0
Total	608	35.2	743	25.3	596	17.9	283	11.5	189	4.9

Note: Recid = recidivism. These recidivism estimates were obtained from survival analysis and are presented for illustrative purposes only. For applied risk assessment reports, readers are encouraged to obtain the most recent recidivism estimates generated from logistic regression analyses (available from www.static99.org).

Table 10. Observed and Predicted Five-Year Recidivism Rates for Static-99 and Static-99R (All Available Cases)

		Observed	Static-99 predicted			Static-99R predicted		
Age	*N*	*n* recid	*n* recid	E/O	95% CI	*n* recid	E/O	95% CI
	Sexual recidivism							
<30	1,405	211	191.4	0.91	0.79-1.04	217.6	1.03	0.90-1.18
30s	1,839	239	209.2	0.88*	0.77-0.99	240.6	1.01	0.89-1.14
40s	1,448	123	142.9	1.16	0.97-1.39	114.7	0.93	0.78-1.11
50s	656	52	58.7	1.13	0.86-1.48	47.2	0.91	0.69-1.19
60s	313	9	27.5	3.06*	1.59-5.87	13.4	1.49	0.77-2.86
70s	79	3	7.2	2.40	0.77-7.44	3.6	1.20	0.38-3.72
50+	1,048	64	93.4	1.46*	1.14-1.86	64.2	1.00	0.78-1.28
	Violent recidivism							
<30	1,135	365	281.8	0.77*	0.70-0.86	316.7	0.87*	0.78-0.96
30s	1,495	347	309.1	0.89*	0.80-0.99	346.5	1.00	0.90-1.11
40s	1,176	165	223.5	1.35*	1.16-1.58	176.3	1.07	0.92-1.24
50s	566	51	98.0	1.92*	1.46-2.53	76.4	1.50*	1.14-1.97
60s	254	10	44.4	4.44*	2.39-8.25	20.8	2.08*	1.12-3.86
70s	64	4	11.2	2.80*	1.05-7.46	5.3	1.32	0.50-3.53
50+	884	65	153.6	2.36*	1.85-3.01	102.5	1.58*	1.24-2.01

Note: Recid = recidivism. CI = confidence interval. Predicted values were obtained from logistic regression analyses using routine versus nonroutine samples as a covariate. Sexual recidivism: for Static-99, $\chi^2 = 27.37$, $df = 5$, $p < .001$; for Static-99R, $\chi^2 = 3.04$, $df = 5$, $p = .694$. Violent recidivism: for Static-99, $\chi^2 = 122.61$, $df = 5$, $p < .001$; for Static-99R, $\chi^2 = 27.29$, $df = 5$, $p < .001$.
*$p < .05$.

E/*O* effect sizes for sexual and violent recidivism rates for Static-99 and Static-99R, separated by age bands and using the entire sample. Static-99 produced *E*/*O* values ranging between 0.88 and 3.06, significantly overestimating recidivism for offenders in their 60s and 70s. For example, for offenders in their 60s, the number of recidivists predicted by Static-99 was three times greater than the observed number of recidivists (*E*/*O* = 3.06, 95% CI of 1.59 to 5.87). Due to the small number of recidivists in the older age groups, a combined category of all offenders aged 50 and older was also examined. Static-99 significantly overestimated recidivism for this combined group (*E*/*O* = 1.46, 95% CI of 1.14 to 1.86). For Static-99R, however, none of the *E*/*O* values were significant, indicating a good fit between observed and predicted sexual recidivism rates (*E*/*O* values ranged between 0.91 and 1.49). The largest *E*/*O* was for offenders in their 60s but this was not statistically significant (*E*/*O* = 1.49, 95% CI of 0.77 to 2.86).

For violent recidivism, both Static-99 and Static-99R yielded predicted recidivism rates per age group that were significantly different than what was observed (for

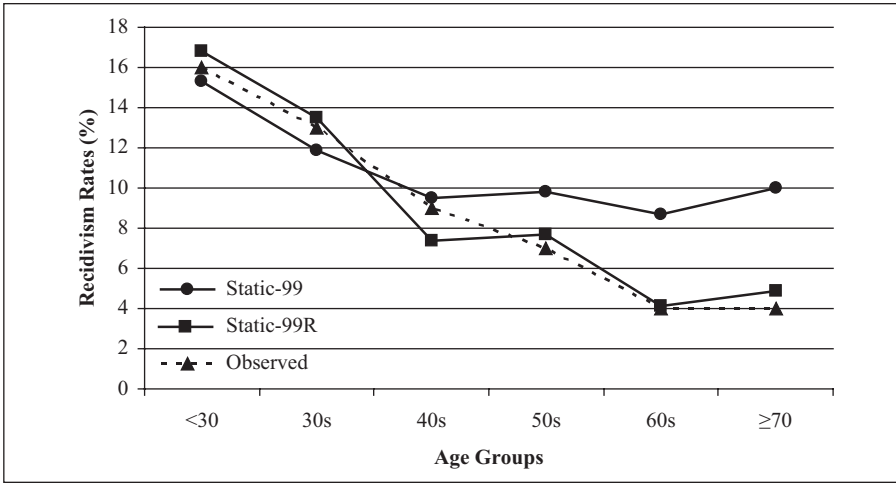

Figure 1. Comparing observed 5-year sexual recidivism rates per age group to recidivism rates predicted from Static-99 and Static-99R, using data from the validation sample

Static-99, $\chi^2 = 122.61$, $df = 5$, $p < .001$; for Static-99R, $\chi^2 = 27.29$, $df = 5$, $p < .001$), although the discordance between observed and predicted was smaller for Static-99R. For the original Static-99, *E*/*O* values ranged between 0.77 and 4.44, and were significant for all age groups. Static-99 significantly underestimated recidivism for all groups less than 40 years of age and overestimated recidivism for all groups ages 40 and above. For Static-99R, *E*/*O* values were generally closer to 1.0, ranging between 0.87 and 2.08. Static-99R significantly underestimated violent recidivism for offenders less than 30 years old, and significantly overestimated recidivism for offenders in their 50s and 60s, as well as for the combined group of offenders above the age of 50.

Given that the revised age weights were created using the development sample, it is possible that the results in Table 10 could reflect overfitting for Static-99R. Table 11 presents the same analyses restricted to the validation sample. Similar results were found, although there were some fluctuations and fewer significant findings, likely due to the reduced sample size. Figure 1 provides a graphical representation of the data in Table 11, demonstrating considerable similarity between observed sexual recidivism rates and those predicted by Static-99R, whereas the original scale overestimated recidivism for offenders age 50 and older.

The differences between Static-2002 and Static-2002R were much smaller (Table 12) and the statistical power was lower due to smaller sample sizes. All offenders above 60 were, therefore, combined into one group. For sexual recidivism, none of the *E*/*O* values were significant for either Static-2002 or Static-2002R (for Static-2002, *E*/*O*s ranged between 0.92 and 2.45; for Static-2002R, *E*/*O*s ranged between 0.90 and 1.75). For violent recidivism, *E*/*O* values for Static-2002 ranged between 0.83 (significantly underestimating recidivism for offenders less than 30) and 3.40 (significantly

Table 11. Observed and Predicted Five-Year Recidivism Rates for Static-99 and Static-99R (Validation Sample)

		Observed	Static-99 predicted			Static-99R predicted		
Age	*N*	*n* recid	*n* recid	E/O	95% CI	*n* recid	E/O	95% CI
	Sexual recidivism							
<30	708	111	108.3	0.98	0.81-1.18	119.1	1.07	0.89-1.29
30s	756	101	89.7	0.89	0.73-1.08	102.0	1.01	0.83-1.23
40s	591	54	56.1	1.04	0.80-1.36	43.8	0.81	0.62-1.06
50s	199	14	19.6	1.40	0.83-2.36	15.3	1.09	0.65-1.84
60s	114	5	10.0	2.00	0.83-4.80	4.6	0.92	0.38-2.21
70s	24	1	2.4	2.40	0.34-17.04	1.2	1.20	0.17-8.52
50+	337	20	32.0	1.60*	1.03-2.48	21.1	1.06	0.68-1.64
	Violent recidivism							
<30	554	186	138.5	0.74*	0.64-0.86	153.5	0.82*	0.71-0.95
30s	566	138	116.5	0.84	0.71-1.00	131.5	0.95	0.81-1.13
40s	455	69	82.8	1.20	0.95-1.52	67.4	0.98	0.77-1.24
50s	169	15	30.4	2.03*	1.22-3.36	24.8	1.65*	1.00-2.74
60s	83	5	14.7	2.94*	1.22-7.06	7.6	1.52	0.63-3.65
70s	22	2	4.2	2.10	0.52-8.40	2.2	1.10	0.28-4.40
50+	274	22	49.3	2.24*	1.48-3.40	34.6	1.57*	1.04-2.39

Note: Recid = recidivism. CI = confidence interval. Predicted values were obtained from logistic regression analyses using routine versus nonroutine samples as a covariate. Sexual recidivism: for Static-99, $\chi^2 = 7.20$, $df = 5$, $p = .206$; for Static-99R, $\chi^2 = 3.43$, $df = 5$, $p = .634$. Violent recidivism: for Static-99, $\chi^2 = 48.24$, $df = 5$, $p < .001$; for Static-99R, $\chi^2 = 15.52$, $df = 5$, $p = .008$.
*$p < .05$.

overestimating recidivism for offenders age 60 and above). For Static-2002R, *E/O* values ranged between 0.82 (significantly underestimating recidivism for offenders less than 30) and 2.52 (significantly overestimating recidivism for offenders age 60 and above). For violent recidivism, both Static-2002 and Static-2002R significantly overestimated recidivism for all offenders age 50 and above.

Discussion

Consistent with previous research (Barbaree & Blanchard, 2008; Hanson, 2006), this study found that the actuarial weights given to age at release in Static-99 and Static-2002 overestimated the risk of older sex offenders. Consequently, new age weights were developed. The revised measures (Static-99R and Static-2002R) were not meaningfully better than the original measures at assessing relative risk for sexual recidivism in the full sample, likely due to the small percentage of offenders influenced by the revised age weights. Static-99R did, however, show substantially better

Table 12. Observed and Predicted Five-Year Recidivism Rates for Static-2002 and Static-2002R (All Available Cases)

		Observed	Static-2002 predicted			Static-2002R predicted		
Age	*N*	*n* recid	*n* recid	E/O	95% CI	*n* recid	E/O	95% CI
	Sexual recidivism							
<30	488	98	95.7	0.98	0.80-1.19	95.1	0.97	0.80-1.18
30s	537	90	83.2	0.92	0.75-1.14	90.6	1.01	0.82-1.24
40s	370	46	46.7	1.02	0.76-1.36	41.4	0.90	0.67-1.20
50s	141	12	14.6	1.22	0.69-2.14	16.0	1.33	0.76-2.35
60+	108	4	9.8	2.45	0.92-6.53	7.0	1.75	0.66-4.66
50+	249	16	24.4	1.52	0.93-2.49	23.0	1.44	0.88-2.35
	Violent recidivism							
<30	442	162	134.0	0.83*	0.71-0.96	133.4	0.82*	0.70-0.96
30s	458	115	111.2	0.97	0.80-1.16	120.6	1.05	0.87-1.26
40s	330	60	68.8	1.15	0.89-1.48	62.3	1.04	0.81-1.34
50s	128	11	22.0	2.00*	1.11-3.61	24.2	2.20*	1.22-3.97
60+	103	5	17.0	3.40*	1.42-8.17	12.6	2.52*	1.05-6.05
50+	231	16	39.0	2.44*	1.49-3.98	36.8	2.30*	1.41-3.75

Note: Recid = recidivism. CI = confidence interval. Predicted values were obtained from logistic regression analyses using routine vs. nonroutine samples as a covariate. Sexual recidivism: for Static-2002, $\chi^2 = 5.03$, $df = 4$, $p = .284$; for Static-2002R, $\chi^2 = 3.19$, $df = 4$, $p = .526$. Violent recidivism: for Static-2002, $\chi^2 = 26.78$, $df = 4$, $p < .001$; for Static-2002R, $\chi^2 = 23.34$, $df = 4$, $p < .001$.
*$p < .05$.

accuracy than Static-99 for predicting absolute recidivism rates. Specifically, the revised scale resulted in less overestimation of recidivism for offenders above 50. Static-2002R also showed improved absolute predictive accuracy compared to the original scale (i.e., there was less overestimation of recidivism for older offenders), but the difference was small, likely because Static-2002 had better age weights than Static-99 to begin with. Despite the small observed difference, we recommend the revised version of Static-2002 as the optimal age weights should logically be the same for both scales.

For violent recidivism, Static-99R demonstrated significantly greater relative predictive accuracy (AUCs) than Static-99. For absolute predictive accuracy, the revision was a clear improvement (i.e., overestimation of recidivism for older offenders was substantially reduced). Age, however, continued to add incrementally to both Static-99R and Static-2002R, and both scales significantly overestimated violent recidivism for the combined group of offenders above the age of 50. For violent recidivism, we conclude that the revisions improved predictive accuracy, but that different weights would have been used if the scale was primarily intended to predict violence.

Age is more strongly related to nonsexual violent recidivism than to sexual recidivism. Consequently, any single set of weights would not be ideal for both outcome

measures. We chose to develop the age weights that best fit the prediction of sexual recidivism because that is the primary outcome for which the scales were designed to predict. That the scales predict violent recidivism with moderate accuracy may be useful for certain evaluators. Evaluators particularly concerned with violent recidivism, however, are advised to use a risk assessment scale specifically designed for that purpose.

The current findings illustrate that the item weightings for actuarial risk tools are unlikely to ever be optimal. Even when actuarial tools are developed on large samples (approximately 1,000), the optimal weights will vary with changes in the offender population. Such fluctuations are inherent in empirically based risk assessment procedures. As noted by Dawes and colleagues (1989), actuarial scales should be revised as more and better research becomes available.

To argue for revisions in actuarial risk measures, however, requires stronger evidence than that previously used to develop the measure. As well, the changes should make substantive differences in interpretation for a nontrivial proportion of offenders. We believe these conditions are met for the proposed age adjustments. The number of offenders in the current sample was larger (N = 8,390) than in the Static-99 development samples (N = 1,208), the samples were more recent, close to 10% of the current sample was above 60 years of age, and the scores for these offenders would change by –3 points (Cohen's d = 1.2). Consequently, for offenders above 60 years of age, the absolute recidivism rate predicted by Static-99R would be roughly half of the recidivism rate predicted by Static-99 (each Static-99 point corresponds to a risk ratio of approximately 1.33; $1.33^{-3} = 0.43$).

The incremental effect of advanced age in the current study is also unlikely to be unique to Static-99 and Static-2002. Given that North American demographic trends suggest the proportion of older offenders will continue to increase, developers of other actuarial scales should further examine the incremental validity of advanced age in their scales.

Although the primary purpose was to improve the actuarial weighting of two widely used risk scales for sex offenders, this study also contributes to the basic research on the relationship between age and crime. This study was not designed to test any of the dominant theories of the age-crime relationship. It does, however, contribute to the basic facts that would need to be explained by any theoretical account.

Most criminological research and theory has focused on explaining why the rates of criminal behavior are much higher among youth compared to adults. One line of explanation has focused on early onset of crime as a marker for persistent, life-course antisociality (Harris & Rice, 2007a; Moffitt, 1993). Another line of research has focused on the social and psychological changes that contribute to desistance from crime during the transition from youth to adulthood (Laub, Nagin, & Sampson, 1998; Maruna, 2001; Serin & Lloyd, 2009; Stouthamer-Loeber, Wei, Loeber, & Masten, 2004). In contrast to both these lines of research, the current study focused on differences between mid and late adulthood. Specifically, sex offenders who are above 60 years old at release had meaningfully lower recidivism rates than those released in

their 30s or 40s, and even those in their 50s. These risk differences remained even after controlling for well-established, static risk markers. The current study also reinforced previous findings that age has a stronger relationship to nonsexual violence than to sexual crime (Hanson & Bussière, 1998).

It is unlikely that a substantial reduction in risk magically occurs on the morning of an offender's 60th birthday. Rather, age predicts recidivism because it is a marker for the underlying propensities for sexual crime. Much of the debate concerning age and sex offender risk assessment concerns the extent to which these propensities are dynamic (i.e., change with age). The findings of Harris and Rice (2007a) suggest that an offender's age is a proxy for enduring antisociality, which is better assessed using age at first offence rather than age at release. Alternately, Barbaree and Blanchard (2008) have suggested a causal effect of aging, whereby sex offenders are less risky as they grow old because of meaningful psychological and physiological changes (e.g., decreased sex drive).

The current study was not designed to directly contribute to this explanatory research, and the findings can support both perspectives. Like other investigators (Harris & Rice, 2007a), we found that older offenders had less extensive criminal histories than younger offenders, supporting age as a marker for enduring antisociality. Consistent with aging effects, however, we also found very low recidivism rates among sex offenders above the age of 60—much lower than would be expected given the offenders' histories of sexual and violent crime. Any explanation of the age–crime relationship, therefore, must address these basic findings. It is likely that both factors (i.e., enduring antisociality and aging effects) contribute to the observed recidivism rates for older offenders. Disentangling the relative contributions of these processes requires further research on the relationship between age, opportunity to offend, and the stability of the psychological and physical characteristics associated with the risk for sexual crime.

In the context of applied risk assessment, the current study extends previous research attempting to develop post-hoc age adjustments to actuarial scales (Barbaree et al., 2007, 2009; Wollert, 2006; Wollert et al., 2010). Although we agree with the previous researchers that age adds incrementally beyond certain static actuarial scales and should be incorporated into risk assessments, we disagree with their specific proposals. Incorporating age directly as an item within the scale is simpler and preferable to making adjustments after the scale has been coded, particularly when (inefficient) age weights are already included in the scales.

Barbaree and colleagues' approach to age adjustments (regression weights) makes use of uncontroversial statistical procedures; however, the estimates provided by this method would need to be validated on new samples to control for overfitting. Overfitting is always a concern for actuarial measures, but it is a particular problem for Barbaree et al.'s (2007, 2009) adjustments given that they were developed on a single sample with an unusually large effect of age.

Wollert's (2006) adjustments, in contrast, require assumptions for which consensus in the scientific community has not been achieved (see critique by Harris & Rice,

2007b). Specifically, we do not believe that likelihood ratios are, or should be, stable features of actuarial risk. Likelihood ratios are defined as the sensitivity/(1–specificity), and they change with the distribution of risk scores in the sample even when the observed recidivism rates per risk score are completely stable across samples. In contrast, our study (and Barbaree et al.'s) directly modeled the effect of age using uncontroversial statistical methods. And in contrast to Barbaree et al. (2007, 2009), our analyses were based on large, diverse samples, which should mitigate the threat of sampling bias. Importantly, the results were robust in cross-validation (no shrinkage was observed).

More recently, Wollert and colleagues (2010) proposed and estimated age-stratified actuarial tables for Static-99. Although age-stratified tables are a plausible solution in principle, their proposal is less efficient and precise than our revisions, with no promise of improved accuracy. Furthermore, their estimates require assumptions about the stability of likelihood ratios (assumptions that we do not share). Separating recidivism tables by age cohort also leads to small sample sizes for offenders above 60, reducing the reliability of the estimates. We believe our revisions are superior because they are based on real data with complete information (i.e., no missing items), specified follow-up periods, and estimated recidivism rates for *each* score.

One limitation of the current study is that, despite accumulating the largest multisite dataset on sexual risk assessment that we are aware of, there were still insufficient cases of older offenders to conduct some potentially informative analyses. For example, we were only able to identify eight offenders age 60 or older who were high risk on Static-99R and we were, therefore, unable to make any meaningful comments on the recidivism rates of this small group of offenders. We were also unable to examine differences in older offenders who are released after serving lengthy prison sentences, compared with older historical offenders (who have potentially been living offence-free in the community for years or decades before being sanctioned) or with older offenders who had committed a recent sex offence.

Implications for Applied Assessment

Even if the reasons for the low sexual recidivism rates of older offenders is not fully understood, the reliable effect of age as a risk marker needs to be considered by evaluators. For current users of Static-99 and Static-2002, we believe there is sufficient evidence to justify switching to the revised age weights (Static-99R/Static-2002R). Given that age at release does not provide incremental predictive validity to Static-99R or Static-2002R for sexual recidivism, use of the revised scales reduces the need for routine post-hoc adjustments for age (such adjustments, however, may still be helpful for predicting violent recidivism). The current data do not preclude the possibility that an evaluator may legitimately assess an older sex offender (e.g., 60+) as high risk. Findings from the current study, however, suggest that these offenders represent exceptional cases.

Appendix

Proportions and Recidivism Rates for Each Risk Category in Original and Revised Scales

	Original scale			Revised scale		
Risk category	*n*	% per category	5-year sex recid	*n*	% per category	5-year sex recid
Static-99						
Low (≤1)	2,380	29.4	4.1	2,836	35.0	4.0
Moderate-low (2,3)	2,685	33.1	6.9	2,209	27.2	7.5
Moderate-high (4,5)	1,830	22.6	14.5	1,837	22.7	13.5
High (6+)	1,211	14.9	25.4	1,224	15.1	26.6
Static-2002						
Low (≤2)	531	20.4	3.6	686	26.3	3.7
Low-moderate (3,4)	693	26.6	6.4	718	27.5	8.3
Moderate (5,6)	714	27.4	13.2	650	24.9	14.5
Moderate-high (7,8)	444	17.0	25.1	369	14.1	26.5
High (9+)	227	8.7	31.8	186	7.1	34.4

Note: Recid = recidivism. Only cases with sexual recidivism data were included. Sexual recidivism was calculated using survival analysis. Evaluators interested in reporting percentiles, relative risk, and recidivism rates associated with Static-99R and Static-2002R should refer to the detailed materials available at www.static99.org (or from the authors).

Authors' Note

The views expressed are those of the authors and not necessarily those of Public Safety Canada or the Wisconsin Department of Health Services.

Acknowledgments

The authors would like to thank Amy Phenix, Andrew J. R. Harris, and Dennis Doren for their contributions throughout this project, as well as the three anonymous reviewers for their helpful feedback on previous versions of this manuscript. They would also like to thank the following researchers for granting us permission to use their data and for being patient with our ongoing questions: Alfred Allan, Tony Beech, Susanne Bengtson, Jacques Bigras, Sasha Boer, Jim Bonta, Sébastien Brouillette-Alarie, Franca Cortoni, Jackie Craissati, Margretta Dwyer, Reinhard Eher, Doug Epperson, Tina Garby, Randolph Grace, Steve Gray, Andy Haag, Leigh Harkins, Grant Harris, Andreas Hill, Steve Johansen, Ray Knight, Niklas Långström, Terry Nicholaichuk, Kevin Nunes, Jean Proulx, Martin Rettenberger, Marnie Rice, Steve Saum, Rebecca Swinburne Romine, Daryl Ternowski, Robin Wilson, and Annie Yessine.

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: Funding for this project was provided in part by the Social Science and Humanities Research Council of Canada and the Ontario Graduate Scholarship Program.

Note

1. Routine correctional samples were defined as relatively random (i.e., unselected) samples from a correctional system, not just from one security level, institution, or treatment program. For example, routine samples could consist of federal offenders, community offenders, offenders serving jail sentences, but not offenders from a specific institution. For more information (including examples) on this variable, see Helmus (2009).

References

References marked with an asterisk indicate studies included in the analyses. Wilson et al. (2007a & 2007b) both have an asterisk, although they were combined into one sample for the current study. The data from Eher et al. (2009) is described in Eher et al. (2008).

Administration on Aging. (2009). *Aging statistics*. Washington, DC: Administration on Aging, United States Department of Health and Human Services. Retrieved from http://www.aoa.gov/AoARoot/Aging_Statistics/index.aspx

*Allan, M., Grace, R. C., Rutherford, B., & Hudson, S. M. (2007). Psychometric assessment of dynamic risk factors for child molesters. *Sexual Abuse: A Journal of Research and Treatment, 19*, 347-367. doi:10.1007/s11194-007-9052-5

Allison, P. D. (1984). *Event history analysis: Regression for longitudinal event data.* Beverly Hills, CA: SAGE.

Archer, R. P., Buffington-Vollum, J. K., Stredny, R. V., & Handel, R. W. (2006). A survey of psychological test use patterns among forensic psychologists. *Journal of Personality Assessment, 87*, 84-94. doi:10.1207/s15327752jpa8701_07

Babchishin, K. M., Hanson, R. K., & Helmus, L. (2011). *The RRASOR, Static-99R, and Static-2002R all add incrementally to the prediction of recidivism among sex offenders* (Corrections Research User Report No. 2011-02). Ottawa, Ontario: Public Safety Canada.

Barbaree, H. E., & Blanchard, R. (2008). Sexual deviance over the lifespan: Reduction in deviant sexual behavior in the aging sex offender. In D. R. Laws & W. T. O'Donohue *Sexual deviance: Theory, assessment, and treatment* (pp. 37-60). New York, NY: Guilford.

Barbaree, H. E., Langton, C. M., & Blanchard, R. (2007). Predicting recidivism in sex offenders using the VRAG and SORAG: The contribution of age-at-release. *International Journal of Forensic Mental Health, 6*, 29-46.

Barbaree, H. E., Langton, C. M., Blanchard, R., & Cantor, J. M. (2009). Aging versus stable enduring traits as explanatory constructs in sex offender recidivism: Partitioning actuarial prediction into conceptually meaningful components. *Criminal Justice and Behavior, 36*, 443-465. doi:10.1177/0093854809332283

*Bartosh, D. L., Garby, T., Lewis, D., & Gray, S. (2003). Differences in the predictive validity of actuarial risk assessments in relation to sex offender type. *International Journal of Offender Therapy & Comparative Criminology, 47*, 422-438. doi:10.1177/0306624X03253850

Beech, A., Friendship, C., Erikson, M., & Hanson, R. K. (2002). The relationship between static and dynamic risk factors and reconviction in a sample of U.K. child abusers. *Sexual Abuse: A Journal of Research and Treatment, 14*, 155-167. doi:10.1177/107906320201400206

*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 & Law, 14*, 85-106. doi:10.1080/10683160701483104

*Bigras, J. (2007). La prédiction de la récidive chez les délinquants sexuels [Prediction of recidivism among sex offenders]. *Dissertations Abstracts International, 68*(09), (UMI No. NR30941).

Boe, R., Nafekh, M., Vuong, B., Sinclair, R., & Cousineau, C. (2003). *The changing profile of the federal inmate population: 1997 and 2002* (Research Report No. R-132). Ottawa, Ontario: Correctional Services Canada.

*Boer, A. (2003). *Evaluating the Static-99 and Static-2002 risk scales using Canadian sexual offenders* (Unpublished master's thesis). University of Leicester, Leicester, United Kingdom.

Bonta, J., Law, M., & Hanson, K. (1998). The prediction of criminal and violent recidivism among mentally disordered offenders: A meta-analysis. *Psychological Bulletin, 123*(2), 123-142. doi:10.1037/0033-2909.123.2.123

*Bonta, J., & Yessine, A. K. (2005). [Recidivism data for 124 released sexual offenders from the offenders identified in *The National Flagging System: Identifying and responding to high-risk, violent offenders* (User Report 2005-04). Ottawa: Public Safety and Emergency Preparedness Canada]. Unpublished raw data.

*Brouillette-Alarie, S., & Proulx, J. (2008, October). *Predictive and convergent validity of phallometric assessment in relation to sexual recidivism risk.* Poster presented at the annual conference for the Association for the Treatment of Sexual Abusers, Atlanta, GA.

*Cortoni, F., & Nunes, K. L. (2007). *Assessing the effectiveness of the National Sexual Offender Program* (Research Report No. R-183; Unpublished report). Correctional Service of Canada.

Craig, L. A., Thornton, D., Beech, A., & Browne, K. D. (2007). The relationship of statistical and psychological risk markers to sexual reconviction in child molesters. *Criminal Justice and Behavior, 34*, 314-329. doi:10.1177/0093854806291416

*Craissati, J., Bierer, K., & South, R. (2008). *What do sex offenders really get up to? Risk prediction, community failure and "sexually risky behaviours" in a nine year follow up study*. Unpublished manuscript. Oxleas NHS Foundation Trust, Bracton Centre, Dartford, United Kingdom.

Dawes, R. M., Faust, D., & Meehl, P. E. (1989). Clinical versus actuarial judgment. *Science, 243*, 1668-1674. doi:10.1126/science.2648573

*Eher, R., Rettenberger, M., Schilling, F., & Pfafflin, F. (2008). Failure of Static-99 and SORAG to predict relevant reoffense categories in relevant sexual offender subtypes: A prospective study. *Sexual Offender Treatment, 8*(1), 1-20. Retrieved from http://www.sexual-offender-treatment.org/sot-1-2008.html

*Eher, R., Rettenberger, M., Schilling, F., & Pfafflin, F. (2009). [Data from sex offenders released from prison in Austria]. Unpublished raw data.

*Epperson, D. L. (2003). *Validation of the MnSOST-R, Static-99, and RRASOR with North Dakota prison and probation samples.* Unpublished Technical Assistance Report, North Dakota Division of Parole and Probation.

Ferguson, G. A. (1976). *Statistical analysis in psychology and education* (4th ed.). New York, NY: McGraw-Hill.

Fleiss, J. L., Levin, B., & Paik, M. C. (2003). *Statistical methods for rates and proportions* (3rd ed.). Hoboken, NJ: Wiley.

Gail, M. H., & Pfeiffer, R. M. (2005). On criteria for evaluating models of absolute risk. *Biostatistics, 6*, 227-239. doi:10.1093/biostatistics/kxi005

Grove, W. M., Zald, D. H., Lebow, B. S., Snitz, B. E., & Nelson, C. (2000). Clinical versus mechanical prediction: A meta-analysis. *Psychological Assessment, 12*, 19-30. doi:10.1037/1040-3590.12.1.19

*Haag, A. M. (2005). [recidivism data from 198 offenders detained until their warrant expiry date. From: Do psychological interventions impact on actuarial measures: An analysis of the predictive validity of the Static-99 and Static-2002 on a re-conviction measure of sexual recidivism. *Dissertations Abstracts International*, *66* (08), 4531B. (UMI No. NR05662)]. Unpublished raw data.

Hanley, J. A., & McNeil, B. J. (1983). A method of comparing the Areas under ROC curves derived from same cases. *Radiology, 148*, 839-843.

Hanson, R. K. (2002). Recidivism and age: Follow-up data from 4,673 sexual offenders. *Journal of Interpersonal Violence, 17*, 1046-1062. doi:10.1177/088626002236659

Hanson, R. K. (2006). Does Static-99 predict recidivism among older sexual offenders? *Sexual Abuse: A Journal of Research and Treatment, 18*, 343-355. doi:10.1007/s11194-006-9027-y

Hanson, R. K. (2008). What statistics should we use to report predictive accuracy. *Crime Scene, 15*(1), 15-17. Retrieved from http://www.cpa.ca/cpasite/userfiles/Documents/Criminal%20Justice/Crime%20Scene%202008-04.pdf

Hanson, R. K., & Bussière, M. T. (1998). Predicting relapse: A meta-analysis of sexual offender recidivism studies. *Journal of Consulting and Clinical Psychology, 66*, 348-362.

*Hanson, R. K., Harris, A. J. R., Scott, T., & Helmus, L. (2007). *Assessing the risk of sexual offenders on community supervision: The Dynamic Supervision Project* (Corrections Research User Report No. 2007-05). Ottawa, Ontario: Public Safety Canada.

Hanson, R. K., Helmus, L., & Thornton, D. (2010). Predicting recidivism among sexual offenders: A multi-site study of Static-2002. *Law and Human Behavior, 34*, 198-211. doi:10.1007/s10979-009-9180-1

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*, 1-21. doi: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*, 119-136. doi:10.1023/A:1005482921333

Hanson, R. K., & Thornton, D. (2003). *Notes on the development of Static-2002.* (Corrections Research User Report No. 2003-01). Ottawa, Ontario: Department of the Solicitor General of Canada.

*Harkins, L., & Beech, A. R. (2007). *Examining the effectiveness of sexual offender treatment using risk band analysis.* Unpublished manuscript. University of Birmingham, Birmingham, United Kingdom.

Harris, A. J. R., Phenix, A., Hanson, R. K., & Thornton, D. (2003). *Static-99 coding rules: Revised 2003*. Ottawa, Ontario: Solicitor General Canada.

Harris, G. T., & Rice, M. E. (2007a). Adjusting actuarial violence risk assessments based on aging or the passage of time. *Criminal Justice and Behavior, 34*, 297-313. doi:10.1177/0093854806293486

Harris, G. T., & Rice, M. E. (2007b). Characterizing the value of actuarial violence risk assessments. *Criminal Justice and Behavior, 34*, 1638-1658. doi:10.1177/0093854807307029

*Harris, G. T., Rice, M. E., Quinsey, V. L., Lalumière, M. L., Boer, D., & Lang, C. (2003). A multi-site comparison of actuarial risk instruments for sex offenders. *Psychological Assessment, 15*, 413-425.

Helmus, L. (2009). *Re-norming Static-99 recidivism estimates: Exploring base rate variability across sex offender samples* (Master's thesis). Retrieved from ProQuest Dissertations and Theses database (UMI No. MR58443) and also from www.static99.org

*Hill, A., Habermann, N., Klusmann, D., Berner, W., & Briken, P. (2008). Criminal recidivism in sexual homicide perpetrators. *International Journal of Offender Therapy and Comparative Criminology, 52*, 5-20. doi:10.1177/0306624X07307450

Hirschi, T., & Gottfredson, M. R. (1983). Age and the explanation of crime. *American Journal of Sociology, 89*, 552-584. doi:10.1086/227905

Hosmer, D. W., & Lemeshow, S. (2000). *Applied logistic regression* (2nd ed.). New York, NY: Wiley.

Humphreys, L. G., & Swets, J. A. (1991). Comparison of predictive validities measured with biserial correlations and ROCs of signal detection theory. *Journal of Applied Psychology, 76,* 316-321.doi:10.1037/0021-9010.76.2.316

Interstate Commission for Adult Offender Supervision. (2007). *Sex offender assessment information survey* (ICAOS Documents No. 4-2007). Lexington, KY: Author.

Jackson, R. L., & Hess, D. T. (2007). Evaluation for civil commitment of sex offenders: A survey of experts. *Sexual Abuse: A Journal of Research and Treatment, 19*, 409-448. doi:10.1007/s11194-007-9062-3

*Johansen, S. H. (2007). Accuracy of predictions of sexual offense recidivism: A comparison of actuarial and clinical methods (UMI No. 3255527). *Dissertations Abstracts International, 68*(03), B.

*Knight, R. A., & Thornton, D. (2007). *Evaluating and improving risk assessment schemes for sexual recidivism: A long-term follow-up of convicted sexual offenders* (Document No. 217618). Submitted to the U.S. Department of Justice.

*Langström, N. (2004). Accuracy of actuarial procedures for assessment of sexual offender recidivism risk may vary across ethnicity. *Sexual Abuse: A Journal of Research and Treatment, 16*, 107-120. doi:10.1177/107906320401600202

Laub, J. H., Nagin, D. S., Sampson, R. J. (1998). Trajectories of change in criminal offending: Good marriages and the desistance process. *American Sociological Review, 63*, 225-238. doi:10.2307/2657324

Maruna, S. (2001). *Making good: How ex-convicts reform and rebuild their lives*. Washington, DC: American Psychological Association.

McGrath, R. J., Cumming, G. F., Burchard, B. L., Zeoli, S., & Ellerby, L. (2010). *Current practices and emerging trends in sexual abuser management: The Safer Society 2009 North American survey.* Brandon, VT: Safer Society Foundation.

Meehl, P. E. (1956). Wanted—A good cook-book. *American Psychologist, 11*, 263-272. doi:10.1037/h0044164

Moffitt, T. E. (1993). Adolescence-limited and life-course-persistent antisocial behavior: A developmental taxonomy. *Psychological Review, 100*, 674-701. doi:10.1037/0033-295X.100.4.674

Mossman, D. (1994). Assessing predictions of violence: Being accurate about accuracy. *Journal of Consulting and Clinical Psychology, 62*, 783-792. doi:10.1037/0022-006X.62.4.783

*Nicholaichuk, T. (2001, November). *The comparison of two standardized risk assessment instruments in a sample of Canadian Aboriginal sexual offenders.* Paper presented at the annual Research and Treatment Conference of the Association for the Treatment of Sexual Abusers, San Antonio, TX.

Olver, M. E., Wong, S. C. P., Nicholaichuk, T., & Gordon, A. (2007). The validity and reliability of the Violence Risk Scale—Sexual Offender Version: Assessing sex offender risk and evaluating therapeutic change. *Psychological Assessment, 19*, 318-329. doi:10.1037/1040-3590.19.3.318

Phenix, A., Doren, D., Helmus, L., Hanson, R. K., & Thornton, D. (2009). *Coding rules for Static-2002*. Ottawa, Ontario: Public Safety Canada. Retrieved from http://www.publicsafety.gc.ca/res/cor/rep/_fl/sttc-2002-eng.pdf

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*, 615-620. doi:10.1007/s10979-005-6832-7

Rockhill, B., Byrne, C., Rosner, B., Louie, M. M., & Colditz, G. (2003). Breast cancer risk prediction with a log-incidence model: Evaluation of accuracy. *Journal of Clinical Epidemiology, 56*, 856-861. doi:10.1016/S0895-4356(03)00124-0

Sampson, R. J., & Laub, J. H. (2003). Life-course desisters? Trajectories of crime among delinquent boys followed to age 70. *Criminology, 41*, 555-592. doi:10.1177/0002716205280075

*Saum, S. (2007). A comparison of an actuarial risk prediction measure (Static-99) and a stable dynamic risk prediction measure (Stable-2000) in making risk predictions for a group of sexual offenders (UMI No. 3255539). *Dissertations Abstracts International, 68*(03), B.

Serin, R. C., & Lloyd, C. (2009). Examining the process of offender change: The transition to crime desistance. *Psychology, Crime & Law, 15*, 347-364.

Soothill, K. L., & Gibbens, T. C. N. (1978). Recidivism of sexual offenders: A re-appraisal. *British Journal of Criminology, 18*, 267-276.

Statistics Canada. (2008). *Report on the demographic situation in Canada: 2005 and 2006* (Catalogue No. 91-209-X). Ottawa, Ontario, Canada: Author.

Stouthamer-Loeber, M., Wei, E., Loeber, R., & Masten, A. S. (2004). Desistance from persistent serious delinquency in the transition to adulthood. *Development and Psychopathology, 16*, 897-918. doi:10.1017/S0954579404040064

*Swinburne Romine, R., Dwyer, S. M., Mathiowetz, C., & Thomas, M. (2008, October). *Thirty years of sex offender specific treatment: A follow-up Study*. Poster presented at the conference for the Association for the Treatment of Sexual Abusers, Atlanta, GA.

*Ternowski, D. R. (2004). Sex offender treatment: An evaluation of the Stave Lake Correctional Centre Program. *Dissertations Abstracts International, 66*(06), 3428B, (UMI No. NR03201).

Thornton, D. (2002). Constructing and testing a framework for dynamic risk assessment. *Sexual Abuse: A Journal of Research and Treatment, 14*, 139-153. doi:10.1023/A:1014620214905

Thornton, D. (2006). Age and sexual recidivism: A variable connection. *Sexual Abuse: A Journal of Research and Treatment, 18*, 123-135. doi:10.1007/s11194-006-9007-2

Turcotte, M., & Schellenberg, G. (2007). *A portrait of seniors in Canada* (Catalogue No. 89-519-XIE). Ottawa, Ontario: Social and Aboriginal Statistics Division, Statistics Canada.

Viallon, V., Ragusa, S., Clavel-Chapelon, F., & Bénichou, J. (2009). How to evaluate the calibration of a disease risk prediction tool. *Statistics in Medicine, 28*, 901-916. doi:10.1002/sim.3517

*Wilson, R. J., Cortoni, F., & Vermani, M. (2007a). *Circles of support and accountability: A national replication of outcome findings* (Report No. R-185). Ottawa, Ontario: Correctional Service of Canada.

*Wilson, R. J., Picheca, J. E., & Prinzo, M. (2007b). Evaluating the effectiveness of professionally-facilitated volunteerism in the community-based management of high-risk sexual offenders: Part two—A comparison of recidivism rates. *Howard Journal, 46*, 327-337. doi:10.1111/j.1468-2311.2007.00480.x

Wollert, R. (2006). Low base rates limit expert certainty when current actuarials are used to identify sexually violent predators. *Psychology, Public Policy, and Law, 12*, 56-85. doi:10.1037/1076-8971.12.1.56

Wollert, R., Cramer, E., Waggoner, J., Skelton, A., & Vess, J. (2010). Recent research (N = 9,305) underscores the importance of using age-stratified actuarial tables in sex offender risk assessments. *Sexual Abuse: A Journal of Research and Treatment, 22*, 471-490. doi:10.1177/1079063210384633