Re-norming Static-99 recidivism estimates: Exploring base rate variability across sex offender samples

A thesis submitted to
the Faculty of Graduate Studies and Research
in Partial Fulfillment of the requirements for the degree

Master of Arts

by

Leslie Helmus

Department of Psychology
Carleton University

September 2009

Abstract

Most replication studies of actuarial risk scales for sex offenders focus on relative predictive accuracy (e.g., ROC), with little attention paid to the stability of absolute recidivism estimates. Research has identified several factors external to actuarial scales that may affect recidivism rates (e.g., recidivism definition, age, dynamic risk factors, treatment, country, setting). Raw data from 29 Static-99 replication studies were combined (*n* = 9,261). Recidivism base rates per Static-99 score were significantly and meaningfully lower in more recent samples compared to the original norms, necessitating new norms. Although Static-99 provided consistent estimates of relative risk, base rates varied across samples. Two variables strongly influenced the base rates: age at release and sample type (defined by the preselection in the samples). A revised version of Static-99 with new age weights was developed (Static-99R), and the implications of base rate variability for actuarial evaluations are discussed.

Acknowledgements

Firstly, this research project would not have been possible without the generousity of researchers in sharing their data (and in responding to my many enquiries). In addition to those whose research is included in the current study, many more researchers offered data that could not be used due to insufficient information or logistical reasons. My sincere gratitude goes to: Alfred Allan, Howard Barbaree, Tony Beech, Susanne Bengtson, Jacques Bigras, Marc Boccaccini, Sasha Boer, Jim Bonta, Sébastien Brouillette-Alarie, Franca Cortoni, Leam Craig, Jackie Craissati, Vivienne de Vogel, Claire Ducro, Margretta Dwyer, Reinhard Eher, Jerome Endrass, Doug Epperson, Tina Garby, Randolph Grace, Steve Gray, Don Grubin, Andy Haag, Neils Habermann, Leigh Harkins, Juliet Harman, Grant Harris, Andreas Hill, Steve Johansen, Ray Knight, Niklas Långström, Calvin Langton, Jan Looman, Bill Marshall, Bob McGrath, John Milton, Steve Montana, Dan Murrie, Terry Nicholaichuk, Kevin Nunes, Jean Proulx, Martin Rettenberger, Marnie Rice, Steve Saum, Cornelis Stadtland, Rebecca Swinburne Romine, Daryl Ternowski, David Thornton, Robin Wilson, and Annie Yessine. The collaborative spirit among researchers in this field never ceases to amaze me.

I would also like to thank my supervisor, Ralph Serin, for his unfailing optimism and support. Thanks to Karl Hanson for sharing some of his knowledge and humour with me. Thanks are also due to David Thornton, Amy Phenix, Andrew Harris, and Dennis Doren for their feedback and guidance throughout this project. Thanks to my committee members (Kevin Nunes, Alfonso Abizaid, and Diana Young) for their helpful comments.

I am also grateful to friends and family. I would like to acknowledge Kelly Babchishin, both for her assistance with some of the tables, graphs, and analyses, as well as for her friendship. Thanks to my husband, Chris, for his support. Thanks also goes to Julie Blais, Leticia Gutierrez, Tania Torrance, and to my parents, Diane and Dirk.

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

Table of Contents

List of Tables

List of Figures

Re-norming Static-99 recidivism estimates: Exploring base rate variability across sex offender samples

Given the severe adverse consequences to victims of sexual offences (Paolucci, Genuis, & Violato, 2001; Resick, 1993), politicians, researchers, and the public have an understandable interest in methods to effectively manage sex offenders and reduce recidivism. Such methods require identifying offenders likely to reoffend. Considerable research has identified factors associated with an increased likelihood of sexual recidivism, such as being young (Hanson, 2006; Hanson & Bussière, 1998; Harris & Hanson, 2004), unmarried (Hanson & Bussière, 1998), having unrelated male child victims, and prior sexual offences (Hanson & Bussière, 1998; Harris & Hanson, 2004).

Risk assessment is a method of evaluating the likelihood of future criminal behaviour by combining multiple risk factors into an overall assessment of recidivism risk (Hanson & Morton-Bourgon, 2009). Risk assessment has become a key activity in our criminal justice system, with profound consequences for public safety and particularly for the offender. Many decisions throughout an offender's progression in the criminal justice system involve risk assessment, including sentencing, security classification, parole decisions, treatment needs, and supervision intensity. Risk assessment is also inherent in two of the three principles of effective correctional treatment (risk/need/responsivity; Andrews & Bonta, 2006; Andrews et al., 1990). Treatment intensity should be directly proportional to the offender's risk (the risk principle), and treatment should target the criminogenic needs of the offender (those needs related to criminal behaviour; the need principle). Adherence to either principle requires a risk assessment.

Given this reliance on risk assessment and the implications for public safety and for the offender, it is important for risk evaluators to be cognizant of the history of risk assessment, its various methods, their empirical support, and the limitations of current research. For actuarial risk assessment scales, such as Static-99, evaluators should also understand the distinction between absolute and relative risk, and how variations in recidivism base rates may affect the absolute recidivism estimates of actuarial tools.

Risk Assessment: History and Various Methods

An early and often-cited milestone in the history of violence risk assessment is the Baxstrom case of 1966, where a Supreme Court ruling resulted in the release of patients judged by clinicians to be criminally insane and highly dangerous (Webster, Douglas, Eaves, & Hart, 1997). Of 98 patients followed for four years, only two were reconvicted for a violent offence (Quinsey, Harris, Rice, & Cormier, 2006), demonstrating that these offenders were not as dangerous as predicted, and leading many to conclude that accurate prediction of violent behaviour was simply not feasible (Hanson, 2005).

Fortunately, in the decades since the Baxstrom case, research on the prediction of violence has flourished with scores of risk assessment instruments developed and validated, showing at least moderate predictive accuracy (Hanson, 2005). Bonta (1996) characterizes this development in three generations. The first generation (applied to the patients in the Baxstrom case) consists of unstructured professional judgement, where a clinician gathers information and forms a subjective risk assessment. The weaknesses of this method are its overreliance on personal discretion and its lack of accountability and replicability (Bonta, 1996).

The second generation of risk assessment relies on instruments that combine primarily static (i.e., historical and unchanging), empirically derived risk factors (Bonta, 1996). In these instruments (commonly referred to as actuarial), items are often scored with either a 0-1 dichotomy (absent-present) or with a specified weighting determined by the strength of the item's relationship to recidivism. The weakness in this generation is that the focus on static factors precludes identification of areas to target in treatment to reduce risk and it cannot reflect positive changes (Bonta, 1996).

The third generation evolved from the second to incorporate criminogenic needs (Bonta, 1996), which are dynamic (i.e., changeable) risk factors that, if changed, can alter the likelihood of reoffending (Andrews et al., 1990). Examples of key criminogenic needs include antisocial personality (e.g., aggression, impulsivity) and antisocial attitudes (Andrews & Bonta, 2006). Third generation scales are therefore sensitive to offender changes and they also tend to have a stronger basis in theories of offending, as well as empirical evidence (Bonta, 1996). Similar to the second generation, these tools are typically actuarial. Recently, Andrews, Bonta, and Wormith (2006) have suggested that a fourth generation of risk assessment has emerged, which provides a comprehensive guide for human service delivery that spans from intake through to case closure.

One area not addressed by Bonta's (1996) description is the status of Structured Professional Judgement (SPJ). SPJ is a method of risk assessment where explicit risk factors (often both static and dynamic) are scored, but the combination of these items into an overall evaluation of risk is left to the judgement of the clinician (Boer, Wilson, Gauthier, & Hart, 1997). Proponents of SPJ argue that clinical judgement should be incorporated in risk assessment because the statistical approach of actuarial scales is not always appropriate in individual cases (Webster et al., 1997). Other researchers, however, have been dismissive of SPJ (Andrews & Bonta, 2006; Bonta, 2002; Quinsey et al., 2006) and classify it as a variation of the first generation of risk assessment (Andrews et al., 2006).

Hanson and Morton-Bourgon (2009) have added to the classification of risk assessment methods by applying a more stringent definition of actuarial scales. Their definition is based on Meehl's (1954) criteria that actuarial scales involve explicit rules to combine pre-specified items into total scores, and empirically-derived estimates of recidivism probability linked to each total score (Hanson & Morton-Bourgon, 2009). Given that several tools satisfying the first criteria of actuarial scales do not include absolute recidivism estimates, Hanson and Morton-Bourgon (2009) made a distinction between actuarial scales (using Meehl's definition) and mechanical scales. Mechanical scales typically contain factors identified based on theory or previous literature reviews, which are combined into a total score based on explicit item weightings, but do not contain a table with recidivism estimates per score.

Which Method Is More Effective?

Most research has focused on comparing the predictive accuracy of actuarial scales and unstructured clinical judgement and has generally favoured actuarial methods (Dawes, Faust, & Meehl, 1989). A large meta-analysis (Grove, Zald, Lebow, Snitz, & Nelson, 2000) examined clinical versus actuarial prediction of human behaviour across various areas of psychology, including medicine ($k = 51$), clinical-personality ($k = 41$), education ($k = 18$), forensic ($k = 10$), financial ($k = 5$), and other ($k = 11$). On average, the results modestly favoured actuarial prediction, with a trend ($p < .07$) for findings in medical and forensic settings to show a stronger advantage for actuarial prediction.

In correctional psychology, a meta-analysis examining violence prediction (Mossman, 1994) favoured actuarial scales (ROC = .71 compared to ROC = .67 for clinical methods). A meta-analysis with mentally disordered offenders found a large difference between actuarial methods and clinical judgement for the prediction of both violent (r of .39 versus .03, respectively) and general recidivism (r of .30 versus .09, respectively; Bonta, Law, & Hanson, 1998).

Hanson and Morton-Bourgon (2009) recently examined four risk assessment methods for sex offenders: actuarial, mechanical, SPJ, and clinical judgement. Individual risk scales were also classified based on whether they were developed for the prediction of sexual, violent, or any recidivism. Results are summarized in Table 1, with outliers excluded. For all three outcomes (sexual, violent, and any recidivism), actuarial and mechanical scales designed to predict sexual recidivism significantly outperformed unstructured clinical judgement. Actuarial and mechanical measures

Table 1

Results from Hanson and Morton-Bourgon's (2009) Meta-Analysis

Form of Risk Assessment	Recidivism Outcome Criteria					
	Sex		Violence		Any	
	d	*k*	*d*	*k*	*d*	*k*
Designed for sexual recidivism						
Empirical-actuarial	.67	81	.51	42	.52	43
Mechanical	.66	29	.40	10	.37	19
Structured professional judgement	.46	6	.31	3	.26	4
Designed for violent recidivism						
Empirical-actuarial	.39	20	.78	15	.74	14
Mechanical	.33	4	.31	3	-	-
Designed for any recidivism						
Empirical-actuarial	.62	9	.79	5	.97	10
Unstructured professional judgement	.42	11	.22	4	.11	9

Note. Categories are left blank when there were less than three findings.

designed to predict sexual recidivism showed similar levels of predictive accuracy when predicting sexual recidivism, but the difference between them increased for the prediction of violent and any recidivism, with actuarial scales demonstrating more of an advantage. The predictive accuracy of structured professional judgement was typically intermediate between actuarial scales and unstructured clinical judgement, although somewhat closer to unstructured judgement. Additionally, each outcome (sexual, violent, or any recidivism) was generally best predicted by risk scales designed for that outcome, with the exception that actuarial risk scales designed to predict general recidivism predicted violence equally as well as actuarial scales designed for violent recidivism.

Actuarial Scales: Absolute Versus Relative Risk

Actuarial risk scales assess two facets of risk: relative and absolute. Relative risk provides information about a particular offender's level of risk relative to other offenders and can be reported in numerous ways, including risk assessment scores (e.g., "this offender is a 3 on Static-99"), nominal risk categories (e.g., "this offender is high risk"), percentiles (e.g., "95% of offenders score higher than this individual"), or relative risk ratios (e.g., "the risk of recidivism for this offender is about ½ the risk of a typical sex offender"; Babchishin & Hanson, 2009).

The accuracy of a risk assessment scale in predicting relative risk can be reported using correlation coefficients, Areas Under the Receiver Operating Characteristic Curve (AUC for ROC), standardized mean differences (Cohen's *d*), or regression coefficients (B_1), and the pros and cons of these statistics have been described elsewhere (Hanson, 2008; Quinsey et al., 2006; Rice & Harris, 2005). The research on predictive accuracy discussed so far in this review has focused on indicators of relative risk. The primary advantage of relative risk information is that it remains fairly consistent across samples (Hanson, Helmus, & Thornton, in press).

Absolute risk, however, refers to the expected probability of recidivism. Although relative risk information is sufficient for most decisions involving the allocation of scarce resources (i.e., treatment and supervision decisions), absolute risk information is required in certain high-stakes evaluations, notably sex offender civil commitment statutes in the United States. Many of these laws require of a determination of whether the offender is more likely than not to reoffend, which has been operationalized as a recidivism estimate of 51% or higher (Doren, 2002).

The probability of recidivism can also be referred to as a base rate. Base rates can be discussed generally (e.g., the recidivism rate for all convicted sex offenders), or for a specific subgroup (e.g., incest offenders). Although the research on predictive accuracy discussed earlier focused on relative risk, base rate information is also important for understanding and contextualizing risk. For example, interpreting the phrase “moderate risk sex offender” requires knowledge of the risk posed by all sex offenders. If the base rate for sexual recidivism is 70%, then the release of a “moderate” risk sex offender invokes concern. If the base rate is 5%, however, then the release of a “moderate” risk sex offender is much less problematic.

Many non-forensic contexts involving the prediction of an event also rely heavily on base rate information. Insurance is the most common example, where base rate statistics (e.g., frequency of car accidents among young males, average life expectancy) are commonly used to determine insurance premiums (Harris, Rice, & Hilton, 2009). Weather forecasts are another example. If asked whether it would snow in Toronto on Canada Day, the forecaster can confidently say “no” because it has never snowed in Toronto in July. If asked, in June, whether it would snow on Christmas Eve, the best response would be a probabilistic estimate based on historical frequencies and current global conditions.

Effective risk communication should incorporate both absolute and relative risk information (Babchishin & Hanson, 2009). Nominal relative risk categories (e.g., low/moderate/high) are interpreted inconsistently (Hilton, Carter, Harris, & Sharpe, 2008; Monahan & Silver, 2003) and with more errors than numerical information (Karelitz & Budescu, 2004). Base rate neglect occurs in risk interpretation when base rate information is overlooked or not fully considered, leading to inaccurate interpretation of relative risk information (Elmore & Gigerenzer, 2005). For example, "high risk" is often interpreted to mean that recidivism is nearly certain, although the base rate for high risk offenders may be fairly low. Generally, people tend to overestimate improbable risk and underestimate more common risks (Evans, Handley, Perham, Over, & Thompson, 2000; Moore, Derry, McQuay, & Paling, 2008). According to the representativeness heuristic, base rate neglect can also occur based on the similarity between the individual being assessed and the referent group (Tversky & Kahneman, 1974). In risk assessment, for example, if an offender appears similar to offenders from which the recidivism estimates were derived, those estimates are interpreted as plausible. If an offender appears different from the referent group, base rate information (even for high risk offenders) may be overlooked and can result in substantial overestimation of risk. Optimal risk assessments should therefore communicate relative risk in the context of absolute risk (i.e., base rates).

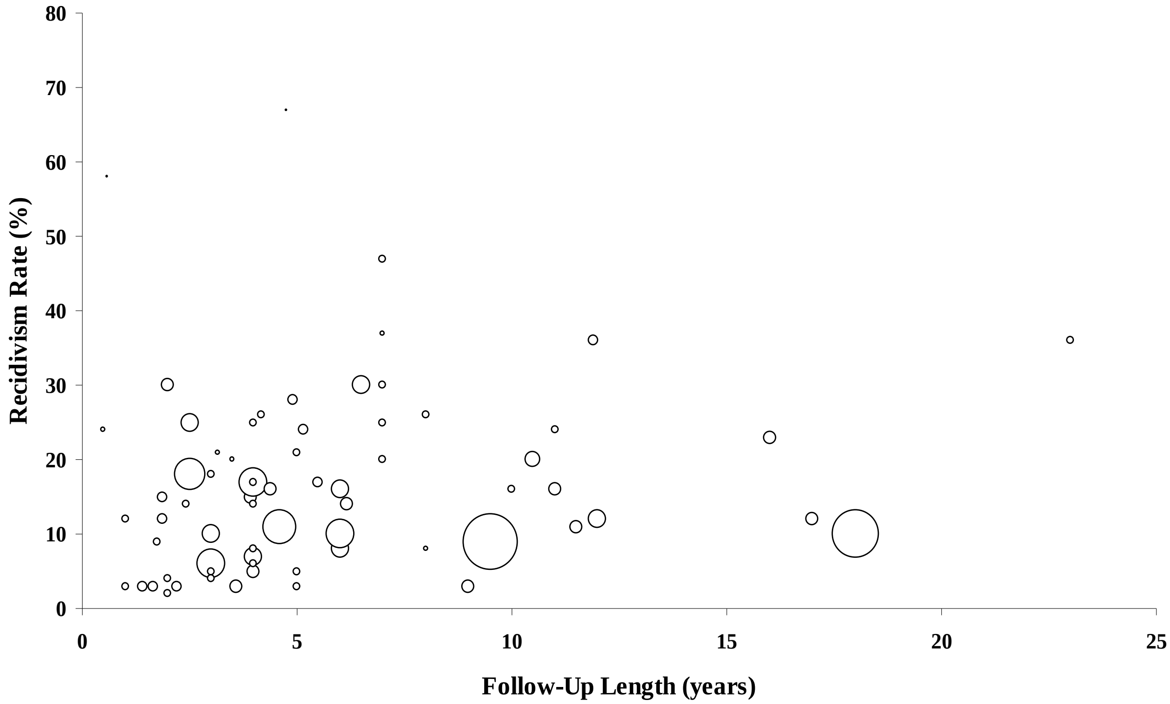
Although base rates are a central part of risk assessment and are routinely reported, relatively little research has examined the stability of base rates for actuarial risk tools for sexual offenders (see Doren, 2004, for an exception). Conventions have yet to be developed concerning the best ways to report this information. Examining this issue should start with an attempt to identify what the base rate for sexual recidivism is and what factors might affect those rates.

What is the Base Rate of Sexual Recidivism?

Some large studies have assessed the overall base rates of sexual recidivism. Hanson and Bussière (1998) found a sexual recidivism rate of 13% among 23,393 offenders ($k = 61$), with an average follow-up period between 4-5 years. Combining 10 samples ($n = 4{,}724$), Harris and Hanson (2004) found sexual recidivism rates of 14%, 20%, and 24% at 5, 10, and 15 years, respectively. Despite the similarity in these aggregate findings at 5 years, the variability across individual studies is substantial. Figure 1 displays sexual recidivism rates from 53 studies in the Hanson and Bussière (1998) meta-analysis, as well as a random sample of 20 newer studies drawn from Hanson and Morton-Bourgon (2009). The recidivism rates are plotted as a function of the follow-up length for the study, with larger bubbles representing larger studies ($N = 35{,}522$).

Interpreting the variability across studies in Figure 1 is difficult because the samples contain a mixture of offenders in terms of their individual risk for recidivism. Actuarial risk scores should therefore explain some of the base rate variability across samples, but variability that persists after controlling for actuarial risk would suggest that there are additional factors not already considered in the actuarial scale that influence base rates. In other words, the recidivism base rates reported with actuarial scores can be viewed as a function of the factors included in the actuarial scale, as well as factors not measured by the scale (plus random error). To the extent that

Figure 1. Sexual recidivism rates (*k* = 73)



variability persists after controlling for actuarial scores, this variability should be noted as a limitation when reporting estimated recidivism rates. This discussion of possible external factors will be divided into three categories:

1) Methodological factors – These factors are part of the design of a study and are typically controlled by the researcher (e.g., how recidivism is measured).

2) Individual-level factors – These factors refer to individual characteristics of the offender. They may or may not be under the

control of the offender (e.g., age is an individual risk factor, but not one that the offender can control).

3) Systems-level factors – These are features of the sample and are typically not under the control of either the researcher or the offender (e.g., the country from which the sample is obtained).

Methodological Factors

Methodological factors are design features of a study, which are typically under the control of the researcher. Broadly, they relate to the nature of the outcome measure (recidivism) and the quality of the actuarial assessment. The outcome should ideally be measured with a lengthy follow-up period using street time to assess an inclusive definition of recidivism gleaned from at least one (but preferably more) reliable recidivism source.

Length of follow-up. Longer follow-up periods increase recidivism base rates because recidivists accumulate over time. This increase, however, is non-linear. Most recidivism occurs within the first few years after release and the longer an offender remains offence-free in the community, the lower their individual probability of recidivism becomes (Harris & Hanson, 2004; A. J. R. Harris, Phenix, Hanson, & Thornton, 2003). Given the relatively low base rate of sexual recidivism, optimal follow-up should be at least five years (Collaborative Outcome Data Committee, 2007b).

Street time. Additionally, using real time or street time in the follow-up calculation can also affect base rates. Real time refers to the actual calendar time that passes between the offender's release from the index offence and when the recidivism information is collected, whereas street time deducts time that the offender does not spend in the community. For example, consider a high-risk sex offender who is released from their index sex offence and for whom recidivism information is collected ten years later. One year into the follow-up period, the offender is arrested for impaired driving causing death and serves eight years in prison. In real time, the sexual recidivism follow-up period is ten years, whereas in street time, the follow-up period is two years. The use of street time should therefore give a more realistic picture of reoffending because it takes into consideration opportunity to reoffend.

Inclusiveness of recidivism definition. More inclusive definitions of sexual recidivism will also logically increase base rates. Examples include using charges as opposed to convictions, or counting any sexually motivated offence (e.g., first-degree murder with a sexual component) as opposed to counting only explicitly sexual offences (e.g., sexual assault). Although Harris and Hanson (2004) did not find higher sexual recidivism rates in samples using charges compared to samples using convictions, studies directly comparing the rate of charges versus convictions within a single sample do find such a difference (Epperson, 2003; Johansen, 2007; Langan, Schmitt, & Durose, 2003; Minnesota Department of Corrections, 2007). For example, of 9,691 sex offenders from 15 states, 5.3% were charged with a new sexual offence within three years, whereas only 3.5% were convicted, suggesting a small but nonetheless meaningful difference (Langan et al., 2003). A similar difference was found in a smaller study (n = 280) with a longer follow-up (average of 7 years), where 6.8% of offenders were rearrested and only 3.9% were reconvicted (Johansen, 2007).

The definition of sexual recidivism does seem to account for some variability in base rates.

Number of recidivism sources. Using multiple sources of recidivism information would result in more accurate data and could also raise the base rate of reoffending. Previous research on two centralized criminal record sources in the United Kingdom (Home Office Offenders Index and the Police National Computer) indicated that both sources contribute unique data on recidivism (Friendship, Thornton, Erikson, & Beech, 2001). Another example comes from the Dynamic Supervision Project (raw data from Hanson, Harris, Scott, & Helmus, 2007), where two recidivism sources (probation/parole officer reports and official national criminal records) revealed a sexual recidivism rate of 6.2%. Additional information obtained from selected provincial criminal records (British Columbia, Manitoba, and Ontario), informal police contacts, and semi-regular perusal of newspapers increased the known recidivism rate to 7.4%.

Quality of the assessment. Methodological factors relating to the quality of the risk assessment can also affect base rates for actuarial scores. Better quality assessments should result in less measurement error, which can affect base rates (in either direction) and reduce the variability in findings. For example, an evaluator who codes an actuarial scale incorrectly or without access to complete information may provide observed base rates for a given score that contains offenders whose “true” score should be higher or lower. Greater confidence in the assessment is expected when scales are scored correctly by conscientious evaluators properly trained, and who have access to complete data. To the extent that any of these elements are missing, base rates may be affected and the relative predictive accuracy may also decrease. For example, Hanson and colleagues (2007) found that conscientious officers (defined as officers who submitted all the data that was requested of them) showed greater predictive accuracy in their Static-99 and Stable-2007 scores (ROCs of .81 and .77, respectively) compared to the complete sample of all officers (ROCs of .74 and .67, respectively). There is also some evidence that studies with greater interrater reliability show significantly larger effect sizes (Hanson & Morton-Bourgon, 2009). Additionally, training by certified trainers increases the validity of risk assessments (Flores, Lowenkamp, Holsinger, & Latessa, 2006) and ongoing training and support are critical for appropriate scoring of actuarial scales (Bonta, Bogue, Crowley, & Motiuk, 2001).

Individual-level Factors

The second category of factors that may affect base rates are individual-level factors not already included in the actuarial scale. It is worth noting that not all external risk factors will add incremental predictive accuracy to an actuarial scale because they may be correlated with other factors already included in the scale.

Dynamic risk factors. Many of the commonly used actuarial risk assessment instruments for sexual offenders focus on static risk factors. Dynamic risk factors are features related to recidivism which can change, and when changed, should alter the likelihood of recidivism (Hanson & Harris, 2000). Some dynamic risk factors for sex offenders identified through meta-analysis include deviant sexual interests, sexual preoccupations, antisocial personality, general self-regulation problems, employment instability, and hostility (Hanson & Morton-Bourgon, 2005). Dynamic risk factors have been found to add incremental predictive validity to static risk factors (Beech, Friendship, Erikson, & Hanson, 2002; Dempster & Hart, 2002) and more recent risk assessment scales have incorporated these factors. Some of the more well-known dynamic risk assessment scales include the Violence Risk Scale – Sex Offender version (Olver, Wong, Nicholaichuk, & Gordon, 2007), the Stable-2007 (Hanson et al., 2007), and the Structured Risk Assessment (Thornton, 2002), all of which have been found to add incremental predictive validity to static actuarial scales.

Treatment. Participation in sexual offender treatment may also contribute to base rates. Several meta-analytic reviews have concluded that sex offender treatment is effective in reducing recidivism (Gallagher, Wilson, Hirschfield, Coggeshall, & MacKenzie, 1999; Hall, 1995; Hanson et al., 2002; Hanson, Bourgon, Helmus, & Hodgson, 2009; Lösel & Schmucker, 2005), although some have argued that there is insufficient evidence to establish treatment effectiveness (Furby, Weinrott, & Blackshaw, 1989; Harris, Rice, & Quinsey, 1998; Kenworthy, Adams, Brooks-Gordon, & Fenton, 2004; Rice & Harris, 2003). One of the largest reviews (Lösel & Schmucker, 2005; Schmucker & Lösel, 2008) examined 69 studies (n = 22,181) and found that the recidivism rate of treated sex offenders was, on average, 6.4 percentage points lower than untreated sex offenders. A more recent meta-analysis (Hanson et al., 2009) examined only studies identified as acceptable study quality according to the Collaborative Outcome Data Committee guidelines (CODC, 2007a, 2007b). Among 22 studies, the unweighted average sexual recidivism rate for treated sex offenders was 10.9%, compared to 19.2% for the comparison offenders.

Although the more recent meta-analytic reviews suggest significant base rate differences among treated versus untreated offenders, they do not answer the question of whether treatment information adds incremental predictive validity to actuarial scores. Evidence for the incremental predictive validity of treatment performance over static actuarial scores has been found in some studies (Marques, Wiederanders, Day, Nelson, & van Ommeren, 2005; Olver et al., 2007), although further research is needed in this area.

Age at release. Age is already considered to some extent in many actuarial scales, (e.g., VRAG, SORAG, Static-99), but age at release has been found to add incremental predictive validity above Static-99, with older offenders showing less sexual recidivism (Hanson, 2006). Additionally, Barbaree, Langton, and Blanchard (2007) found that VRAG and SORAG scores were correlated with age at release, and once the shared variance was removed, age added significantly to the prediction of recidivism.

Systems-Level Factors

Systems-level factors influence recidivism base rates but are typically not under the control of either the offender or the researcher. The most obvious examples of systems-level factors are the jurisdiction and setting of a sample.

Country. International variability in official crime rates has been documented (Krohn, 2001; Maffei & Merzagora Betsos, 2007; Rushton, 1995; United Nations, 2007), including differences in the official rates of sexual offences (Kutchinsky, 1991; United Nations, 2007). Crime rates are higher in countries with greater urbanization and industrialization (Krohn, 2001). Some of this variation could be due to differences in laws as well as prosecution practices; crime rates could be higher when the number of activities defined as illegal increases and when there are more resources (e.g., financial, political) to prosecute certain offences. For example, in the early 1980s in Canada, drastic broadening of the legislation defining sexual offences was accompanied by large increases in officially reported sexual crimes (Brennan & Taylor-Butts, 2008).

Religiosity may also play a part in the variability of international crime rates. Data from 13 industrial nations found an inverse relationship between overall criminality and religiosity, but a reverse trend was found for sexual offences, with higher rates in countries reporting greater levels of religiosity (Ellis & Peterson, 1996). This could be due to an increased tolerance of sexual violence among fundamentalist religions emphasizing the subservience of women (Raj, Silverman, Wingood, & Diclemente, 1999).

Jurisdictional base rate variation could also reflect differences in the quality of criminal record-keeping (Marenin, 1997), with poor or unreliable records resulting in artificially lower base rates. Both Canada and the United Kingdom contain centralized national criminal records. In Canada, these records are maintained by the Royal Canadian Mounted Police in the Canadian Police Information Center (CPIC). CPIC records have certain disadvantages, however, because charges not resulting in convictions are inconsistently recorded, and information is submitted after a disposition is made (e.g., conviction, dismissal, acquittal, stay of proceedings), which can result in substantial delays between the commission of an offence and its appearance on the criminal records. Additionally, records are purged over time due to pardons and inactivity, further reducing their reliability (Hanson & Nicholaichuk, 2000). The United Kingdom has two centralized criminal record sources: the Offenders Index (OI) and the Police National Computer (PNC). The disadvantage of these sources is that they exclude Scotland and Northern Ireland, and the OI includes only standard list offences, which tend to be the more serious types of offences (Friendship et al., 2001).

Although criminal records in Canada and the United Kingdom are far from perfect, they are nonetheless centralized and relatively comprehensive, whereas the United States lacks a comprehensive and centralized criminal record database. Each state maintains its own criminal records and the FBI maintains a separate database, but both sources are known to have incomplete, inaccurate, and ambiguous data (Laudon, 1986).

Setting. In addition to differences across countries, there may be differences across settings and sample types. Most studies of sex offenders do not use a complete (i.e., random) sample of offenders. Instead, samples are often preselected based on certain characteristics. For example, some studies are from a particular treatment setting, or a setting where offenders are referred for assessment or other services (e.g., psychiatric assessment), or from a particular institution (e.g., a maximum security prison). In other studies, the sample is defined by certain conditions the offender meets (e.g., their sentence type or other special measures they are subject to). An important empirical question involves the extent to which offenders preselected under some of these conditions would be expected to vary in their risk for recidivism from truly random, unselected samples of offenders.

It is likely that some of the selection processes described above would select offenders on the basis of factors already included, at least to some extent, in the actuarial scales (e.g., prior sex offences). It is also likely that factors external to actuarial scales would affect these preselection processes. Some of these external factors would presumably be related to risk for reoffending (e.g., treatment need, institutional behaviour, treatment performance), while others may not be (e.g., offence severity, treatment availability, publicity surrounding a case). Normative data have shown that general offenders from institutional samples show consistently higher actuarial risk scores than offenders serving community sentences (Andrews, Bonta, & Wormith, 2004), showing that even crude forms of preselection (sentence type) do seem to distinguish offenders with different risk profiles.

It is plausible that studies of sex offenders over-sample from settings where offenders are preselected to be higher risk. Analyzing a research question requires a sample with sufficient information on the variables of interest, and offenders who are higher risk may disproportionately find themselves in settings or circumstances likely to collect the types of information of interest to researchers. For example, in Canada, specialized psychological assessments are more common for offenders being considered for a Dangerous Offender or Long-Term Offender designation. To the extent that correctional systems invest more time and resources (resulting in more data) on offenders they are more concerned about, then sex offender studies will likely include samples that are systematically higher risk than random (unselected) samples.

Examining different ways offenders are preselected and the extent to which this preselection contributes to recidivism rates may increase our understanding of how some variables external to actuarials are influencing base rates. Preliminary publications from the current project found a large effect for sample type, with samples preselected to be higher risk showing more than double the sexual recidivism rate per actuarial score (using Static-99) as routine samples from the Correctional Service of Canada (Helmus, Hanson, & Thornton, 2009). Recommendations to present separate recidivism estimates for each group and determine which group the offender being evaluated resembles has been meet with criticism, suggesting that applying this distinction in routine practice is problematic, particularly outside Canada (for a summary of these criticisms, see Abbott, 2009). As such, further efforts are needed to define a sample type variable with strong empirical support, but also greater practical utility.

Time period. Cohort effects can also contribute to base rate variability across samples. Crimes rates peaked in the early 1990s and have been generally declining since then. This trend has been found for both violent and property offences in Canada (Public Safety Canada, 2008; Mishra & Lalumière, 2009b) and the United States (Federal Bureau of Investigation, 2007; Mishra & Lalumière, 2009b), using both official crime data as well as victimization surveys (Bureau of Justice Statistics, 2006). Sexual offences appear to be no exception (for a review, see Mishra & Lalumière, 2009a). Declines have been observed in the rates of forcible rape in the U.S. (Federal Bureau of Investigation, 2007), sexual assault in Canada (Mishra & Lalumière, 2009a), clergy sexual abuse (Terry, 2008), and child sexual abuse measured both by substantiated cases as well as victimization surveys (for a summary, see Finkelhor & Jones, 2006; Jones & Finkelhor, 2006). Recent data from Minnesota (*n* = 1,782; Minnesota Department of Corrections, 2007) show a dramatic decline in three-year rates of sexual rearrest, reconviction, and reincarceration. Interestingly, most risky behaviours typically correlated with criminal behaviour (e.g., accidents, suicide, risky sexual behaviour such as unprotected sex, dropping out of school), have also shown similar declines (Mishra & Lalumière, 2009b). Given these overall trends and their apparent universality in both Canada and the United States, it is expected that recidivism base rates will show similar changes over time.

Detection rates. Recidivism rates can also be affected by detection rates because reporting an offence is a necessary precondition for counting an offender as a recidivist. Sexual offences typically have the highest levels of non-reporting (Besserer & Trainor, 2000), with estimated rates between 78-84% (Besserer & Trainor, 2000; Kilpatrick, Edwards, & Seymour, 1992). These startlingly high levels of underreporting, however, do not necessarily mean that the majority of sexual recidivists are not caught. If an offender reoffends with multiple victims (which is not uncommon), only one victim needs to report the offence for the offender to be counted as a recidivist. It is also possible that offenders previously charged with sexual offences are more likely to be caught in the future, as more supervision may be in place (e.g., monitoring by police or probation officers) and potential victims may be more inclined to report a new offence. Underreporting will, therefore, have some impact on recidivism base rates, but the exact relationship is difficult to estimate.

Correctional philosophy. Another factor that could affect recidivism base rates includes features inherent in the correctional system from which the sample is obtained. Considerable research among general offenders has demonstrated that treatment adhering to the principles of effective correctional practice (matching treatment intensity to risk, targeting criminogenic factors, and delivering treatment appropriate to the abilities and learning style of offenders) produces significant reductions in recidivism (Andrews & Bonta, 2006). Similar results have also been obtained with sex offender treatment (Hanson et al., 2009). Additionally, considerable research has found that punitive approaches (e.g., longer sentences) and poor quality treatment do not reduce reoffending, and may be associated with slight increases in reoffending (Andrews & Bonta, 2006; Smith, Goggin, & Gendreau, 2002). Given this body of research, jurisdictions with more punitive correctional systems (e.g., longer and harsher sentences, less treatment) may show higher base rates than systems with more rehabilitative approaches.

Community supervision. The existence and restrictiveness of community supervision is another feature of correctional systems that may have some impact on base rates. Previous meta-analytic research on community supervision found a significant reduction in general recidivism, although the magnitude of this difference was sufficiently small as to be of little value (Bonta, Rugge, Scott, Bourgon, & Yessine, 2008). Further analyses of these studies, however, revealed that community supervision that was delivered in adherence to the principles of effective correctional treatment showed larger effect sizes (Simpson, 2009).

Although community supervision delivered with appropriate treatment may reduce reoffending, highly restrictive supervision may have the opposite effect. There is some evidence that more intensive supervision with home visits results in higher detection rates for new offences (Stalans, Seng, Yarnold, Lavery, & Swartz, 2001). Additionally, intensive supervision programs with high revocation rates for technical breaches may artificially suppress recidivism in the short term by reducing the opportunity to commit new offences. This effect can be mitigated somewhat by analyses using street time as opposed to real time.

Summary of Factors External to Actuarials

Table 2 presents a summary of the methodological, individual, and systems factors discussed above. Though not an exhaustive list, it highlights the number and diversity of factors that could affect recidivism base rates, although assessing their impact is likely to be complicated by their intercorrelations. For example, cohort differences may be caused by a variety of other factors, including an aging population, as well as changes in correctional environments over time (e.g., providing better quality treatment). Another example is that differences across countries could be related to differing correctional philosophies as well as the quality of criminal records. A further conceptual complication is that some systems-level factors may be a proxy for individual-level factors. For example, offenders are often referred to various settings on the basis of individual characteristics (e.g., risk, treatment needs), so distinctions based on setting could reflect individual-level differences in offenders referred to those settings.

Table 2

Summary of Factors That May Affect Recidivism Base Rates Per Actuarial Score

Methodological Factors	Individual-Level Factors	Systems-Level Factors
Length of follow-up	Dynamic risk factors	Country
Street time	Treatment	Setting
Inclusiveness of recidivism definition	Age at release	Time period
Number of recidivism sources		Detection rates
Quality of the risk assessment		Correctional philosophy
		Community supervision

Given the large number of factors external to actuarial scales for which direct or indirect evidence suggests they might impact base rates, it may be naïve to expect consistency across samples in the recidivism rates estimated for each actuarial score. If the evidence suggests that base rates vary significantly based on external factors, this variability should be reported as a limitation in the prediction of absolute recidivism rates. Overcoming this limitation would require further research on how to incorporate that variability to improve the actuarial prediction of recidivism. The form that this improvement takes will depend on what types of variables are found to influence base rates. For example, if methodological factors are more important than individual-level factors, then it would be important to standardize research methods (e.g., the measurement of recidivism) to improve prediction. If individual-level factors are more important, however, improving prediction could involve revising current actuarial scales to incorporate the unmeasured individual characteristics.

This type of research is difficult to conduct because testing the variables discussed above requires recidivism data and actuarial scores from a large number of offenders from a variety of studies using different samples and methodologies. Static-99 is the actuarial risk scale for sex offenders with the most validation studies (Hanson & Morton-Bourgon, 2009), and therefore has the most potential for further examination of these factors.

Development of Static-99

Static-99 is an empirically derived actuarial risk assessment tool designed to predict sexual recidivism among adult male offenders with a charge or conviction for a sexual offence (Hanson & Thornton, 2000; see also www.static99.org). It was developed by combining two actuarial risk assessment instruments for sex offenders: the RRASOR (Hanson, 1997), which focused more on sexual deviance, and the SACJ-Min (Hanson & Thornton, 2000), which included sexual deviance but also had a strong focus on criminal history. In the development study, Hanson and Thornton (2000) combined three Canadian samples and one U.K. sample (n = 1,208), and found that Static-99 predicted sexual and violent recidivism significantly better than either the RRASOR or the SACJ-Min.

Static-99 contains 10 items (see Appendix A), coded as either a 0 or a 1, except for prior sexual offences which is scored as 0, 1, 2, or 3. Total scores (obtained by summing all the items) can range from 0-12. Based on an offender's total score, they are placed in one of four risk categories: low (0-1), moderate-low (2-3), moderate-high (4-5), and high (6+). Static-99 can be scored based on commonly available criminal history information without interviewing the offender (A. J. R. Harris et al., 2003). With training, it can also be coded by virtually anyone, including parole and probation officers, psychologists, treatment providers, and even police officers (A. J. R. Harris et al., 2003).

Static-99 contains sexual and violent recidivism estimates per score at 5, 10, and 15 years (with all offenders scoring 6 and above grouped in one category, due to small numbers; A. J. R. Harris et al., 2003). These recidivism estimates were developed based on three of the four development samples (n = 1,086) and consisted primarily of offenders released in the 1960s, 1970s, and 1980s (Hanson & Thornton, 2000).

Use and Replication of Static-99

Since its development, Static-99 has quickly become the most commonly used actuarial risk tool in Canada and the U.S. for treatment planning (McGrath, Cumming, & Burchard, 2003), community supervision (Interstate Commission for Adult Offender Supervision, 2007), and civil commitment evaluations (Jackson & Hess, 2007). As well, it is used in jurisdictions as diverse as Sweden, Belgium, Israel, Singapore, and Japan (personal communication with Karl Hanson, February 5, 2007).

Not only is it one of the most commonly used, but Static-99 is also the most researched risk assessment instrument for sex offenders. In Hanson and Morton-Bourgon's (2009) meta-analysis, there were 63 replications of Static-99, which was substantially more than any other instrument (aside from its predecessor, the

RRASOR, all other instruments had 12 or less). Numerous replications have demonstrated robust relative predictive accuracy across many settings and for many types of offenders. For example, in addition to numerous replications in prison settings (Beggs & Grace, 2005; Brown, 2003; Craig, Beech, & Browne, 2006; Epperson, 2003; Friendship, Mann, & Beech, 2003; G. T. Harris et al., 2003; Hood, Shute, Feilzer, & Wilcox, 2002; McGrath, Hoke, Livingston, & Cumming, 2001; Langström, 2004; Ternowski, 2004; Thornton, 2002), moderate to large effect sizes have also been found in psychiatric settings (de Vogel, de Ruiter, van Beek, & Mead, 2004; Ducro & Pham, 2006; G. T. Harris et al., 2003; Nunes, Wexler, Firestone, & Bradford, 2003), and community settings (Craissati, Webb, & Keen, 2005; Epperson, 2003; Hanson et al., 2007).

Other replications have demonstrated predictive accuracy among both child molesters and rapists (Bartosh, Garby, Lewis, & Gray, 2003; Ducro & Pham, 2006), with mixed results found for non-contact offenders (Bartosh et al., 2003; Helmus & Hanson, 2007). Replications with developmentally delayed offenders (Tough, 2001) and sexually abusive priests (Montana & Thompson, 2003) have found small and moderate effect sizes, respectively. New analyses of the 63 replication studies included in Hanson and Morton-Bourgon (2009) have found that Static-99 shows moderate to large predictive accuracy in Canada, the United States, the United Kingdom, continental Europe, Australia, and New Zealand, with particularly large effect sizes in the United Kingdom and Australia/New Zealand (Helmus, Hanson, & Morton-Bourgon, in press).

Compared to other actuarial risk scales, Static-99 predicts sexual recidivism reasonably well (*d* = .67), although not notably better than other actuarial tools (see Table 3 for summary of data from Hanson & Morton-Bourgon, 2009). Its popularity, however, demonstrates the widespread demand for cost-effective risk tools applicable to a wide range of sexual offenders.

Summary and Purpose of Current Study

The purpose of the current study is to investigate whether the original recidivism norms for Static-99 are applicable to current samples, and whether recidivism rates per Static-99 score remain stable across samples. Although the predictive accuracy of Static-99 is not notably stronger than most other actuarial scales, it was chosen for this project because the large number of validations provides a rich source of data for cross-sample comparisons. The current study will assess the variability in recidivism base rates across samples, after controlling for Static-99 scores, and explore moderator variables that may account for any observed variability.

Table 3

Hanson & Morton-Bourgon's (2009) Results for Individual Actuarial Risk Scales

Measure	Mean *d*	*k*
MnSOST-R	.76	12
Static-2002	.70	8
Risk Matrix – 2000 sex	.67	10
Static-99	**.67**	**63**
SORAG	.62	12
RRASOR	.60	34
VRAG	.52	8
SIR	.51	5
LSI-R (& variants)	.45	3
SACJ-Min	.42	6

Method

Measures

Static-99. As previously discussed, Static-99 is an empirically derived actuarial risk assessment tool designed to predict sexual and violent recidivism in adult male sex offenders (Hanson & Thornton, 2000; see also www.static99.org). It has ten items and the total score (ranging from 0 – 12) can be used to place offenders in one of four risk categories (low, moderate-low, moderate-high, and high; A. J. R. Harris et al., 2003).

Samples

To evaluate the absolute recidivism estimates of Static-99, raw data from Static-99 replication studies were obtained. To be included, a dataset required sufficient recidivism information (i.e., fixed follow-up outcomes) to conduct logistic regression analyses and complete information for all Static-99 items (samples that approximated some items were excluded). A total of 29 samples were obtained. Prior to merging, each dataset was cleaned to check for internal inconsistencies (e.g., miscalculation of total scores or item scores contradicted by other information in the dataset). Identified errors were corrected if possible; otherwise, the case was deleted. Cases were also deleted under the following circumstances: missing follow-up information, any missing Static-99 item 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 (the age and gender exclusionary criteria are specified in the Static-99 coding rules).

Tables 4 through 8 provide basic descriptive information about the studies included. The total sample consists of 9,261 sex offenders, all male. From Table 4, sample sizes for individual studies varied between 66 and 1,278, with a median sample size of 233. Twelve samples were from Canada, six were from the United States, four were from the United Kingdom, and there was one each from Denmark, Austria, Holland, Sweden, Switzerland, Germany, and New Zealand. Most offenders were from correctional settings ($k = 22$), while 7 samples included offenders from mental health settings or mixed mental health and correctional settings. Of the 17 studies that could be classified in terms of their treatment status, 9 samples were mostly treated (more than 75% of the sample), whereas 6 were mixed in their treatment exposure, and only 1 sample was mostly untreated (less than 25%; one additional sample was divided into two subgroups, one of which was mostly treated and the other was mixed).

From Table 5, age at release was available in 24 samples ($M = 40$, $SD = 12$, with within-sample means varying between 33 and 46). Year of release information was available for over 95% of the sample ($n = 8{,}924$), with offenders released between 1957-2007. Despite the large range, most samples were relatively recent, with 81% of offenders released in 1990 or later. Inter-rater reliability was reported in 11 studies and was consistently high, with correlations ranging from .86 to .92, and Intra-Class Correlation coefficients ranging from .84 to .95.

Table 6 summarizes information on sample-level moderator variables regarding recidivism measurement and the quality of the Static-99 assessment. Thirteen samples used charges as the recidivism criteria, 14 used convictions, and 1

Table 4

Descriptive Information

Study	Sample Size	Mean Static-99 Score (SD)	Country	Setting	Type of Sample	Mostly Treated
Allan et al. (2007)	493	2.2 (2.0)	New Zealand	Corrections	Prison treatment	Yes
Bartosh et al. (2003)	186	3.5 (2.4)	U.S.	Corrections	Routine correctional	
Bengtson (2008)	311	3.8 (2.1)	Denmark	Mixed	Preselected risk/psychiatric	
Bigras (2007)	483	2.7 (2.0)	Canada	Corrections	Routine CSC	Mixed
Boer (2003)	299	3.3 (2.3)	Canada	Corrections	Routine CSC	
Bonta & Yessine (2005)	155	5.2 (1.9)	Canada	Corrections	Preselected high risk	Mixed
Brouillette-Alarie & Proulx (2008)	228	3.9 (2.2)	Canada	Corrections	Routine CSC & community	
Cortoni & Nunes (2007)	73	2.6 (1.8)	Canada	Corrections	CSC low and moderate risk	Yes
Craig et al. (2006)	66	2.4 (1.7)	U.K.	Corrections	Referred for assessment	
Craissati et al. (2008)	209	2.4 (2.0)	U.K.	Corrections	Routine community	Mixed
de Vogel et al. (2004)	121	6.0 (1.7)	Holland	Mental Health	Indeterminate sentence	Yes
Eher et al. (2008)	706	2.7 (2.0)	Austria	Corrections	Routine European prison	
Endrass et al. (2009)	95	3.5 (1.7)	Switzerland	Corrections	Routine European prison	
Epperson (2003)	178	2.7 (2.2)	U.S.	Corrections	Routine correctional	
Haag (2005)	198	3.9 (2.0)	Canada	Corrections	Preselected high risk	Mixed
Hanson et al. (2007)	702	2.9 (2.0)	Canada	Corrections	Routine CSC & community	
Harkins & Beech (2007)	198	2.8 (2.2)	U.K.	Corrections	Prison & community treatment	Yes
Harris et al. (2003)	368	3.4 (2.4)	Canada	Mixed	Preselected risk/psychiatric	Mostly & Mixed
Hill et al. (2008)	86	4.9 (1.8)	Germany	Mixed	Sexual homicide perpetrators	
Johansen (2007)	273	3.0 (2.0)	U.S.	Corrections	Treatment sample	Yes
Knight & Thornton (2007)	466	4.5 (2.2)	U.S.	Mixed	Preselected high risk	
Långström (2004)	1,278	2.4 (2.0)	Sweden	Corrections	Routine European prison	No
Langton (2003)	354	3.6 (2.1)	Canada	Corrections	CSC moderate treatment	Yes
Milton (2003)	118	4.9 (2.0)	U.K.	Mental Health	Psychiatric	Mixed
Nicholaichuk (2001)	281	4.6 (2.0)	Canada	Mental Health	High intensity treatment	Yes
Saum (2007)	175	2.0 (1.6)	U.S.	Corrections	Community	Yes
Swinburne Romine et al. (2008)	681	1.9 (2.1)	U.S.	Corrections	Community treatment	Mixed
Ternowski (2004)	247	2.1 (1.9)	Canada	Corrections	Treatment	Yes
Wilson et al. (2007a & b)	233	5.5 (2.0)	Canada	Corrections	Preselected high risk	

Table 5

Additional Descriptive Information

Study	Sample Size	Age at Release Mean (SD)	Release Period	Median Release Year	Inter-Rater Reliability Statistic	Inter-Rater Reliability
Allan et al. (2007)	493	42 (12)	1990-2000	1994		
Bartosh et al. (2003)	186	38 (12)	1996	1996	*r*	.90
Bengtson (2008)	311	33 (10)	1978-1995	1986	ICC	.94
Bigras (2007)	483	43 (12)	1995-2004	1999		
Boer (2003)	299	41 (13)	1976-1994	1990		
Bonta & Yessine (2005)	155	39 (9)	1992-2004	1998		
Brouillette-Alarie & Proulx (2008)	228	36 (10)	1979-2006	1996		
Cortoni & Nunes (2007)	73	42 (12)	2001-2004	2003		
Craig et al. (2006)	66					
Craissati et al. (2008)	209	38 (12)	1992-2005	1998		
de Vogel et al. (2004)	121		1977-2000	1990	ICC	.92
Eher et al. (2008)	706	41 (12)	2000-2005	2003	ICC	.90
Endrass et al. (2009)	95					
Epperson (2003)	178	37 (13)	1989-1998	1995		
Haag (2005)	198	37 (10)	1995	1995	*r*	.92
Hanson et al. (2007)	702	42 (13)	2001-2005	2002	ICC	.91
Harkins & Beech (2007)	198	43 (12)	1994-1998	1995		
Harris et al. (2003)	368	36 (11)	1968-1997	1989	ICC	.87
Hill et al. (2008)	86	39 (11)	1971-2002	1989	ICC	.84
Johansen (2007)	273	38 (11)	1994-2000	1996		
Knight & Thornton (2007)	466	36 (11)	1957-1986	1970	*r*	.86
Långström (2004)	1,278	41 (12)	1993-1997	1995		
Langton (2003)	354		1990-2001	1995	ICC	.88
Milton (2003)	118		1978-1998	1986		
Nicholaichuk (2001)	281	35 (9)	1983-1998	1992		
Saum (2007)	175	46 (12)			ICC	.95
Swinburne Romine et al. (2008)	681	38 (12)	1977-2007	1988		
Ternowski (2004)	247	44 (13)	1994-1998	1996		
Wilson et al. (2007a & b)	233	42 (11)	1994 -2007	2002		
Overall	9,261	40 (12)	1957-2007	1995		

Table 6

Sample-level Moderator Variables

Study	Sample Size	Recidivism Criteria	# of Recidivism Sources	Used National Records for Recidivism?	Used Street Time?	Cited Coding Rules?
Allan et al. (2007)	493	Charges	1	Yes	No	Yes
Bartosh et al. (2003)	186	Charges	1	Yes	No	Yes
Bengtson (2008)	311	Charges	1	Yes	No	Yes
Bigras (2007)	483	Charges	1	Yes	No	No
Boer (2003)	299	Conviction	1	Yes	No	Yes
Bonta & Yessine (2005)	155	Conviction	2	Yes	No	Yes
Brouillette-Alarie & Proulx (2008)	228	Conviction	2	Yes	No	Yes
Cortoni & Nunes (2007)	73	Charges	1	Yes	No	Yes
Craig et al. (2006)	66	Conviction	1	Yes	No	Yes
Craissati et al. (2008)	209	Conviction	4	Yes	No	No
de Vogel et al. (2004)	121	Conviction	1	Yes	No	No
Eher et al. (2008)	706	Conviction	1	Yes	No	Yes
Endrass et al. (2009)	95	Reincarceration	1	Yes	No	Yes
Epperson (2003)	178	Charges	1	No	No	No
Haag (2005)	198	Conviction	1	Yes	No	Yes
Hanson et al. (2007)	702	Charges	4	Yes	Yes	Yes
Harkins & Beech (2007)	198	Convictions	2	Yes	No	No
Harris et al. (2003)	368	Charges	1	Yes	No	No
Hill et al. (2008)	86	Conviction	1	Yes	Yes	No
Johansen (2007)	273	Charges	2	No	No	Yes
Knight & Thornton (2007)	466	Charges	4	Yes	Yes	Yes
Långström (2004)	1,278	Conviction	1	Yes	No	Yes
Langton (2003)	354	Conviction	1	Yes	No	No
Milton (2003)	118	Conviction	2	Yes	No	No
Nicholaichuk (2001)	281	Conviction	1	Yes	No	No
Saum (2007)	175	Charges	1	No	No	No
Swinburne Romine et al. (2008)	681	Conviction	1	No	No	Yes
Ternowski (2004)	247	Charges	1	No	Yes*	Yes
Wilson et al. (2007a & b)	233	Charges	1	Yes	No	No

*Street time was used in the time-to-recidivism data for survival analysis and Cox regression, but was unavailable for analyses using fixed follow-up periods.

sample used reincarceration for a minimum of 10 months as the recidivism criteria (Endrass et al., 2009); this sample was analyzed with those using reconviction as their outcome. Most studies used one source for collecting recidivism information (*k* = 21), whereas 5 studies used two sources and 3 studies used four sources. Twenty-four of the studies used national criminal records for at least one of their recidivism sources; the remaining studies used local (e.g., state or provincial) records.

Table 7 lists individual-level (i.e., case-level) moderator variables and indicates which samples had information on those variables. Because studies varied in what information was provided and how it was coded, some variables were coded in multiple ways, with the goal that at least one version of a variable would have sufficient cases for analysis. These variables (as well as the exact sample sizes for each level of the moderator variable) will be discussed further in the results section.

Offender type information was available in 15 samples. Non-contact offenders (e.g., exhibitionists, voyeurs) and mixed offenders (those with both adult and child victims) were identified inconsistently and in small numbers. This study will therefore focus on the distinction between rapists and child molesters (insufficient information was available to distinguish between extrafamilial and intrafamilial child molesters). The current literature provides no clear consensus on how to classify offenders as rapists or child molesters. Although the simplest definitions would be based on criminal code designations, these would fluctuate across jurisdictions and over time. Most samples did not specify what definition was used to classify offenders as rapists or child molesters. Variability across samples in definitions is therefore expected. When sufficient victim information was available, however, the

Table 7

Individual-level Moderator Variables: Were Cases Available?

Study	Treatment Status	Started Treatment	Completed Treatment	Offender Type	Canadian Federal vs. Provincial	Aboriginal	NonWhite	NonWhite (excluding Aboriginal)
Allan et al. (2007)	Yes	Yes	Yes	Yes				
Bigras (2007)				Yes	Yes			
Boer (2003)				Yes	Yes	Yes	Yes	Yes
Bonta & Yessine (2005)		Yes			Yes	Yes	Yes	Yes
Brouillette-Alarie & Proulx (2008)				Yes	Yes	Yes	Yes	Yes
Cortoni & Nunes (2007)		Yes			Yes			
Craissati et al. (2008)	Yes	Yes	Yes	Yes				
de Vogel et al. (2004)		Yes						
Eher et al. (2008)				Yes				
Haag (2005)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Hanson et al. (2007)				Yes	Yes	Yes		
Harkins & Beech (2007)		Yes						
Harris et al. (2003)				Yes	Yes			
Hill et al. (2008)						Yes	Yes	Yes
Johansen (2007)		Yes						
Knight & Thornton (2007)		Yes		Yes			Yes	Yes
Langton (2003)				Yes	Yes			
Nicholaichuk (2001)		Yes		Yes	Yes	Yes		
Saum (2007)		Yes	Yes	Yes				
Swinburne Romine et al. (2008)	Yes	Yes	Yes	Yes		Yes	Yes	Yes
Ternowski (2004)	Yes	Yes	Yes		Yes			
Wilson et al. (2007a & b)		Yes		Yes	Yes			

Information was missing for: Bartosh et al., 2003; Bengtson, 2008; Craig et al., 2006; Endrass et al., 2009; Epperson, 2003; Långström, 2004; Milton, 2003.

following definition was applied: offenders with victims less than 14 were considered child molesters, and offenders with adult victims (age 18+) were classified as rapists. For offenders with teen victims between the ages of 14-17 inclusive, they were classified as child molesters if their victims were related and rapists if the victims were unrelated. Offenders with both adult and child victims were classified based on their predominant victim choice (if possible) or were designated as mixed offenders.

Table 8 presents average Static-99 scores as well as sexual (n = 8,893) and violent (n = 7,627) recidivism outcome data for each sample. The average Static-99 score for the total sample was in the moderate-low category (M = 3.1, SD = 2.2), although averages varied substantially across studies, ranging from 1.9 (Swinburne Romine, Dwyer, Mathiowetz, & Thomas, 2008) to 6.0 (de Vogel et al., 2004). The average length of follow-up was 8.1 years (SD = 5.0), with sample-level average follow-ups varying between 3.4 years (Hanson et al., 2007) and 16.8 years (Swinburne Romine et al., 2008). The recidivism rates presented are the observed rates (not controlling for follow-up time) and rates from fixed 5 and 10 year follow-up periods. Note that these data do not control for Static-99 scores. All but one sample contained information on sexual recidivism (G. T. Harris et al., 2003, contained information on violent recidivism only). Combining all 29 samples (N = 9,261), approximately 30% of the data is lost when analyses are restricted to fixed 5-year follow-up periods (n = 6,544 remaining), and 70% of the data is lost when 10-year fixed follow-ups are examined (n = 2,766 remaining). More fluctuation across samples is therefore observed at the 10 year follow-up. The observed sexual

Table 8

Recidivism Information

Study	Mean Static-99 Score (*SD*)	Mean Follow-Up in Years (*SD*)	Overall Sample Size	Cases available at 5 years	Cases available at 10 years	Sexual Recidivism (%) Overall	Sexual Recidivism (%) 5 year	Sexual Recidivism (%) 10 year	Violent Recidivism (%) Overall	Violent Recidivism (%) 5 year	Violent Recidivism (%) 10 year
Allan et al. (2007)	2.2 (2.0)	5.8 (2.9)	493	299	25	9.5	11.7	20.0	16.4	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)	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)	4.6 (1.9)	483	207		6.2	9.2		15.3	22.2	
Boer (2003)	3.3 (2.3)	13.3 (2.1)	299	299	295	8.7	3.7	7.8	23.4	14.0	21.0
Bonta & Yessine (2005)	5.2 (1.9)	5.7 (2.3)	155	101	3	17.4	18.8	0.0	39.4	45.5	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	
Craig et al. (2006)	2.4 (1.7)	10.0 (0.0)	66		66	13.6		13.6	19.7		19.7
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
de Vogel et al. (2004)	6.0 (1.7)	11.6 (5.9)	121	100	71	40.5	26.0	38.0			
Eher et al. (2008)	2.7 (2.0)	3.9 (1.1)	706	151		4.0	2.0		14.7	11.9	
Endrass et al. (2009)	3.5 (1.7)	5.0 (0.0)	95	95		8.4	8.4		10.5	10.5	
Epperson (2003)	2.7 (2.2)	7.9 (2.4)	178	151	36	14.0	10.6	22.2			
Haag (2005)	3.9 (2.0)	7.0 (0.0)	198	198		25.3	19.7				
Hanson et al. (2007)	2.9 (2.0)	3.4 (1.0)	702	31		8.1	0.0		16.4	3.2	
Harkins & Beech (2007)	2.8 (2.2)	10.4 (1.1)	198	198	129	14.1	9.6	16.3	20.7	13.6	23.3
Harris et al. (2003)	3.4 (2.4)	7.6 (4.3)	368	259	119				48.1	38.6	52.1
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)	8.6 (2.6)	466	433	353	26.2	24.7	30.0	36.9	32.5	41.3

Table continues.

Table 8 continued.

Study	Mean Static-99 Score (*SD*)	Mean Follow-Up in Years (*SD*)	Overall Sample Size	Cases available at 5 years	Cases available at 10 years	Sexual Recidivism (%)			Violent Recidivism (%)		
						Overall	5 year	10 year	Overall	5 year	10 year
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
Langton (2003)	3.6 (2.1)	6.4 (3.0)	354	226	47	11.0	10.2	12.8	24.3	22.6	36.2
Milton (2003)	4.9 (2.0)	10.2 (5.0)	118	93	68	16.1	16.1	25.0	28.0	26.9	38.2
Nicholaichuk (2001)	4.6 (2.0)	6.4 (4.0)	281	168	59	18.5	22.6	25.4			
Saum (2007)	2.0 (1.6)	5.0 (0.0)	175	175		35.4	31.4				
Swinburne Romine et al. (2008)	1.9 (2.1)	16.8 (7.8)	681	570	543	13.8	8.4	11.2			
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)	233	104	16	10.3	11.5	6.3	25.8	31.7	43.8
Overall	3.1 (2.2)	8.1 (5.0)	9,261	6,544	2,647	12.9	11.4	17.2	23.9	20.9	32.3

recidivism rate for all cases was 12.9%, with sample-level rates varying between 0% (Cortoni & Nunes, 2007) and 40.5% (de Vogel et al., 2004). At 5 years, the observed sexual recidivism rate was 11.4%, with samples ranging from 0% (Cortoni & Nunes, 2007) to 31.4% (Saum, 2007). At 10 years, the observed sexual recidivism rate was 17.2%, with samples ranging from 0% (Bonta & Yessine, 2005) to 38% (de Vogel et al., 2004). The observed violent recidivism rate for all cases was 23.9%, with samples ranging from 8.2% (Cortoni & Nunes, 2007) to 51.8% (Bengtson, 2008). At 5 years, the observed violent recidivism rate was 20.9%, with samples ranging from 3.2% (Hanson et al., 2007) to 45.5% (Bonta & Yessine, 2005). At 10 years, the observed violent recidivism rate was 32.3%, with samples ranging from 17.7% (Johansen, 2007) to 52.1% (G. T. Harris et al., 2003). Readers should note that the 10 year rates can be lower than the 5 year rates because different offenders were included in the 5 and 10 year samples.

Further descriptions of the individual samples will be reported briefly below. For more detailed information, readers are encouraged to obtain the original studies.

1) Allan, Grace, Rutherford, & Hudson, 2007. The study sample consists of child molesters who completed the prison-based Kia Marama sex offender treatment program in Christchurch, New Zealand. Recidivism information was collected from the New Zealand Department of Corrections.

2) Bartosh, Garby, Lewis, & Gray, 2003. The study sample consists of sex offenders released from the Arizona Department of Corrections and subject to registration and notification. The Static-99 was scored from file information and recidivism was coded from FBI records. Interrater reliability was reported (*r* = .90), although the number of cases coded by multiple raters is unknown.

3) Bengtson, 2008. The study sample consists of sex offenders who received a pre-trial forensic psychiatric evaluation in Denmark. Such evaluations were typically conducted for offenders suspected of mental disorder or mental retardation, offenders deemed high risk by the courts, those accused of serious offences, and those for whom an indefinite sentence was being considered. The Static-99 was coded from file information and criminal records. Recidivism information was obtained from the Danish Central Crime Register, and interrater reliability was assessed by having two raters code 20 cases (*ICC* = .94).

4) Bigras, 2007. The original sample contained 94% of all sexual offenders receiving a federal sentence (two or more years) in Quebec between 1995 and 2000 (6% refused participation in the research or were unable to provide consent). Assessment information was extracted from file data and interviews with offenders. Recidivism data was collected using CPIC records.

5) Boer, 2003. The study sample consists of all male federal offenders serving a sentence for a sexual offence in British Columbia whose Warrant Expiry Date (WED; the end of their sentence) was between January 1990 and May 1994. Many offenders are granted conditional release prior to their WED; thus, offenders in this sample were released as early as 1976. Recidivism information was collected using CPIC records. Category B sexual offences (see A. J. R. Harris et al., 2003) were excluded from the definition of sexual recidivism.

6) Bonta & Yessine, 2005. The original sample consisted of three subgroups of Canadian offenders: 1) offenders flagged as potential Dangerous Offenders (subject to indeterminate sentence) by the National Flagging System, 2) offenders designated as Dangerous Offenders, and 3) offenders who committed a violent reoffence after being detained until their Warrant Expiry Date. Only offenders in the first group (flagged offenders), however, had Static-99 scores available. For these offenders, Static-99 was coded from file information and recidivism was coded from CPIC records and Offender Management System (OMS) records from the Correctional Service of Canada (CSC). The definition of sexual recidivism excluded prostitution offences, indecent phone calls, and possession of child pornography. Given the low frequency of these offences, it is expected that this restricted definition would have minimal impact on the results.

In some cases the offender's "current" offence (i.e., the offence that precipitated the flag) was non-sexual but there was a prior sexual offence on record. Their most recent sex offence was used as the index sex offence for Static-99 scoring purposes (as per the coding rules), but these cases are somewhat unique because the offenders spent time in the community after their index sex offence but before the recidivism follow-up period began. To retain a sample of offenders who were serving a sentence for a sexual offence or who had a recent sex offence on file, offenders with more than two years between their index sex offence and the current offence for which they were flagged were deleted (n = 22).

7) Brouillette-Alarie & Proulx, 2008. The study sample consists of offenders treated or assessed at a maximum security psychiatric facility in Quebec. A sample from this facility was used in the original development samples of Static-99, but any overlapping cases were removed in the present study. Recidivism information was collected from CPIC records and Montreal courthouse records.

8) Cortoni & Nunes, 2007. The study sample consists of Canadian federal offenders who received the low or moderate intensity National Sexual Offender Treatment Program (implemented in 2000). The original study also used a comparison group of federal offenders under community supervision in 1991, but these offenders were excluded from the present study due to insufficient follow-up information for logistic regression analyses. Recidivism information was collected from CPIC records.

9) Craig, Beech, & Browne, 2006. The study sample consists of sex offenders referred for assessment to a Regional (Medium) Secure Unit of a forensic psychiatric facility in the United Kingdom. Although the setting was a forensic psychiatric facility, it provided assessment resources and outpatient treatment services for nonmentally disordered offenders on community supervision. The original study contained a second sample already included in another dataset obtained for this study (Harkins & Beech, 2007). The second sample was therefore removed from this dataset to avoid overlap. Recidivism information (coded from the Home Office Offenders Index) was available only for a ten year fixed follow-up period and therefore could not be used for survival analyses, Cox regression analyses, and logistic regression analyses at 5 years.

10) Craissati, Bierer, & South, 2008. The study sample consists of all contact sex offenders on probation in two boroughs in South East London during the study period. The Static-99 was coded from file information and recidivism data was collected from four sources: the Police National Computer, the Violent and Sex Offenders Register, the Multiple Criminal Remote Access, and the EApps database.

11) de Vogel, de Ruiter, van Beek, & Mead, 2004. The study sample consists of sex offenders in the Netherlands on an indeterminate TBS order (they were adjudicated as having diminished criminal responsibility due to severe psychopathology) in a forensic psychiatric hospital. Static-99 was coded from file information and recidivism was coded from the Judicial Documentation Register. Interrater reliability was assessed by having three raters code 30 cases (*ICC* = .92).

12) Eher, Rettenberger, Schilling, & Pfafflin, 2009. The study sample consists of sex offenders released from prison in Austria (see Eher, Rettenberger, Schilling, & Pfafflin, 2008). Interrater reliability was assessed by having four raters code 27 cases (*ICC* = .90). Recidivism information was collected from the Federal Department of the Interior.

13) Endrass, Urbaniok, Held, Vetter, & Rosseger, 2009. The study sample consists of sex offenders in the Canton of Zurich (Switzerland) who had been sentenced to a minimum prison sentence of 10 months. Static-99 was coded from file information and recidivism data was coded from penal records. Information on time to recidivism was unavailable, so this sample could not be used for survival analysis or Cox regression.

14) Epperson, 2003. The study sample consists of sex offenders in North Dakota who were either incarcerated or on probation. Recidivism information was collected from North Dakota state records.

15) Haag, 2005. The original study sample included all male Canadian federal sex offenders whose Warrant Expiry Date was in 1995, although 75% of offenders were released prior to their WED. Follow-up information was collected for 7 years after the WED. Because recidivism information was not recorded for the time period after release but before the WED, offenders who were released more than 30 days in advance of their WED were deleted, effectively reducing the sample to offenders who were detained until Warrant Expiry. Under Canadian legislation, offenders are to be automatically released after serving two thirds of their sentence. In some cases, however, CSC will make an application to have the offender detained until Warrant Expiry if the parole board is satisfied that if released, the offender poses a significant risk of committing a serious offence before their sentence expires. Recidivism information was collected from CPIC records. Interrater reliability was assessed by having 66 cases from the original sample coded by the main researcher and a CSC psychologist (r = .92).

16) Hanson, Harris, Scott, & Helmus, 2007. This prospective study followed offenders on community supervision between 2001-2005 in Canada, Alaska, and Iowa, although only Canadian offenders were used in the current study. Static-99 was coded by community supervision officers and sent to the project staff, and interrater reliability was examined through file review of 88 cases coded by the officers (*ICC* = .91). Recidivism information was collected from CPIC records, supervising officers, provincial records, and informal police contacts (additionally, one recidivist was identified in a newspaper article).

17) Harkins & Beech, 2007. This sample included offenders from three other studies in the United Kingdom. The first sample consisted of offenders from the mid to late 1990s in Her Majesty's Prison Service's Sex Offender Treatment Programme. The second sample was drawn from a community treatment program in the early 1990s, and the third sample included offenders who underwent community treatment in the late 1990s in West Midlands. The intensity of the treatment varied and treatment dropouts were retained in the sample. Recidivism data was collected from the Home Office Offenders Index and Police National Computer.

18) Harris, Rice, Quinsey, Lalumière, Boer, & Lang, 2003. The study sample consists of Canadian sex offenders from the following sources: 1) offenders assessed in the Sexual Behaviour Lab in Penetanguishene, Ontario, who were either incarcerated after the assessment or released to the community; 2) offenders from the Regional Treatment Centre, located in a maximum security federal institution in Ontario; and 3) offenders from the Regional Psychiatric Centre, a maximum security federal institution in B.C. Static-99 was coded from file information, with some modifications from the coding rules (personal communication between Karl Hanson and Grant Harris, October, 2004). Interrater reliability was assessed by having two raters code 10 cases. Recidivism information was coded from CPIC records. This is the only study included which did not have sexual recidivism; only violent recidivism information was available.

19) Hill, Haberman, Klusmann, Berner, & Briken, 2008. The study sample consists of offenders who committed a sexual homicide. Static-99 was coded from file information and recidivism was coded from German federal criminal records.

20) Johansen, 2007. The study sample consists of participants from a prison-based sex offender treatment program in Washington State. Static-99 was coded from file information and recidivism was coded from the Office of the Administrators of the Courts and the Department of Corrections Offender Based Tracking System.

21) Knight & Thornton, 2007. This study followed offenders who were either assessed or treated at the Massachusetts Treatment Center (MTC; a treatment center for sexually dangerous persons) between 1959 and 1984. Static-99 was retrospectively coded from file data by raters who were blind to recidivism status. Recidivism information was obtained from four sources: Massachusetts Board of Probation records, Massachusetts Parole Board records, MTC Authorized Absence Program records, and FBI records. Interrater reliability was evaluated by having 232 Static-99 cases coded by two raters (r = .87). In the original dataset, total scores were averaged across raters as opposed to generating a consensus rating. For this project, a consensus score was identified wherever possible. For example, given inconsistent coding concerning age at release, the offender's birth date and release date were used to obtain the correct score. Also, if one rater's score did not make sense given other information in the dataset, then the other rater's score was used. When consensus ratings could not be generated, the rating was determined by flipping a coin.

22) Långström, 2004. The study sample consists of sex offenders released from prison in Sweden. The Static-99 was coded from file information and recidivism was coded from the National Council for Crime Prevention.

23) Langton, 2003. This study followed Canadian federal sex offenders offered treatment at Warkworth Sexual Behaviour Clinic (WSBC) between 1989 and

2001. WSBC offered moderate intensity treatment for sexual offenders within a medium security institution in Ontario. Information to code Static-99 was extracted from Warkworth's clinical files, OMS records, and CPIC records. Recidivism information was coded from CPIC records. Interrater reliability was assessed by having two researchers code 25 cases for Static-99 (r = .88).

24) Milton, 2003. The study sample consists of mentally disordered sex offenders referred to treatment within one of three high security hospitals in the United Kingdom. Static-99 was coded from file information and recidivism was coded from the Home Office Offenders Index and the mental health files of the hospital.

25) Nicholaichuk, 2001. The study sample consists of sex offenders treated at the Clearwater sex offender treatment program, located in a federal maximum-security forensic mental health facility in Saskatchewan. Recidivism information was coded from CPIC records.

26) Saum, 2007. The study sample consists of sex offenders treated by North Dakota's Department of Human Services. Static-99 was scored from file information and recidivism was coded from records maintained by the North Dakota Department of Corrections. Interrater reliability was assessed by having two raters code 20 cases (ICC = .95).

27) Swinburne Romine, Dwyer, Mathiowetz, & Thomas, 2008. The study sample consists of offenders assessed or treated at an outpatient treatment program at the University of Minnesota. Recidivism information was coded from the Bureau of Criminal Apprehension records.

28) Ternowski, 2004. The study sample consists of sex offenders referred to the Stave Lake Correctional Centre Program, located in a provincial institution in B.C. Recidivism information was coded from provincial correctional records maintained by the Minister of the Attorney General of British Columbia.

29) Wilson and colleagues (2007a & b). The study sample consists of Canadian offenders combined from two previous studies: Wilson, Cortoni, and Vermani (2007a), and Wilson, Picheca, and Prinzo (2007b). Both studies consist of high-risk sex offenders who were detained in prison until their Warrant Expiry Date (the end of their sentence). In both studies, half of the offenders participated in Circles of Support and Accountability, while another (matched) group of sex offenders did not. Although the two studies had separate samples, they were combined into one dataset because both samples were selected in the same way and the basic descriptive information was the same for both studies.

Original Static-99 Development Samples

To compare the recidivism estimates from the current samples to the original Static-99 recidivism norms, the raw dataset used to produce the original Static-99 recidivism estimates was obtained. The original estimates were calculated using three of the four Static-99 development samples: two samples from Canada (Institut Philippe Pinel and Millbrook Correctional Centre) and one from the United Kingdom (offenders released from Her Majesty's Prison Service). These samples are described in more detail elsewhere (Hanson & Thornton, 2000).

Moderator Variables

Because the data were obtained from pre-existing datasets, analyses are limited to information already coded. A summary of the moderator variables available for testing is provided below, along with the hypotheses about their effect on base rate variability across samples, after controlling for Static-99.

1) Use of charges or convictions as recidivism definition. The hypothesis was that samples using charges would have higher recidivism rates than those using convictions.

2) Number of recidivism sources. Higher base rates were expected among samples using multiple recidivism sources.

3) Use of national criminal records. The hypothesis was that recidivism rates would be higher in samples using centralized national criminal records as their source for recidivism information. The rationale for this hypothesis was that more complete sources of data would increase base rates.

4) Street time. The hypothesis was that studies using street time in their follow-up calculations would have higher recidivism base rates than studies using real time.

5) Did the study cite the Static-99 coding rules? This variable was used to assess the quality of the risk assessment. Although it is a relatively crude proxy for assessment quality, it is plausible to assume that authors who cite the coding rules are more likely to be aware that official coding rules exist, and are therefore more likely to adhere to these coding rules. A study was coded positive on this variable if they cited either the 2003 coding rules (A. J. R. Harris et al., 2003) or earlier/unofficial coding rules if the study pre-dated the 2003 version. Additionally, one study (Knight

& Thornton, 2007) was coded as a “yes” on this variable even though the coding rules were not cited because one of the authors was co-developer of Static-99. The hypothesis was that the quality of the assessment will influence the variability of base rates as opposed to the direction. Specifically, relatively high variability among studies not citing the coding rules was expected.

6) Canadian jurisdiction. No hypotheses were made about base rate differences among offenders in the Canadian federal system versus offenders in the provincial system. On the one hand, offenders with sentences of two or more years may have lengthier criminal histories than provincial offenders, which would increase their risk for reoffence. Alternately, however, federal offenders have increased opportunities to participate in more intensive treatment programs, which may reduce their risk for reoffence.

7) Offender type. Previous research has found that rapists and child molesters show similar sexual recidivism rates (Harris & Hanson, 2004), but violent recidivism rates are higher among rapists (Hanson et al., in press). As such, the hypothesis was that rapists and child molesters would show similar sexual recidivism base rates.

8) Country. Given that crime rates are higher in the United States compared to Canada (Federal Bureau of Investigation, 2007; Mishra & Lalumière, 2009b; Public Safety Canada, 2008), incarceration rates are higher in the U.S. compared to the rest of the western world (Public Safety Canada, 2008), and that the U.S. is known for more punitive correctional policies compared to Canada (Mauer, 2006), the hypothesis was that base rates would be higher in the United States as compared to Canada. Given the relative similarity in crime rates between Canada and many

European countries (United Nations, 2007), Canada and Europe were hypothesized to show similar base rates.

9) Age at release. The hypothesis was that base rates per Static-99 score would be inversely affected by age at release (lower recidivism among older offenders).

10) Release year. This variable was used to examine possible cohort effects. Given declining crime rates since the early 1990s, the hypothesis was that there would be a negative linear relationship between year of release and base rates (i.e., lower recidivism rates would be found among more recent releases).

11) Race. Although race has not been found to be a significant predictor of sexual recidivism (Hanson & Bussière, 1998), it was included as a moderator variable because some research among general offenders has found higher recidivism base rates among Aboriginals in particular (Bonta, Laprairie, & Wallace-Capretta, 1997; Broadhurst & Loh, 2005). No hypotheses were made regarding this moderator variable.

12) Treatment. The hypothesis was that offenders who completed treatment will show lower recidivism rates per Static-99 score than offenders who did not complete treatment, and that treatment dropouts would have the highest recidivism rates.

13) Sample type. The premise of this variable was that some study samples include all sex offenders within a correctional system, while other samples examine restricted groups of offenders (e.g., offenders within a specific treatment program, or offenders who have been given indeterminate sentences). The hypothesis was that samples representative of general populations of sex offenders would show lower base rates than samples preselected on risk relevant characteristics.

Overview of Analyses

Logistic regression analysis at fixed 5 year and 10 year follow-ups were used to examine both relative and absolute risk properties of Static-99 (Hosmer & Lemeshow, 2000). 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), logistic regression estimates two regression coefficients (B_0 and B_1). B_1 (often referred to as a slope) is an estimate of predictive accuracy, or the average change in recidivism rates for adjacent scores. In other words, B_1 estimates the average increase in recidivism (expressed as a log odds ratio) associated with each one-unit increase in Static-99. B_0 (often referred to as an intercept) is a measure of the recidivism base rate for the sample (technically, the log odds of the predicted recidivism rate for offenders with a Static-99 score of zero). For ease of interpretation, the B_0 values can be transformed into percentages.

For samples in which there were no recidivists, logistic regression coefficients could not be computed. Rather than deleting these low base rate samples, the recidivism base rate (p) was estimated as $1/4n$ (i.e., Bartlett's adjustment, see Eisenhart, 1947, §4.3; Cohen, 1988, p. 183). In this formula, n refers to the total number with follow-up information, not just those with a particular score, because if there were any recidivists in the sample, B_0 would not equal zero. For meta-analysis, the proportions were transformed into log odds ratios (the same metric as the logistic regression coefficients), with an estimated variance of $1/(np(1-p))$ (see Fleiss, Levin, & Paik, 2003, §2.6).

It is important to note that the B_0 is a base rate estimate for Static-99 scores of zero. Re-centering Static-99 can produce B_0s that examine base rates for any score. Examining base rates at multiple Static-99 scores obtains a more accurate picture of variability because base rate differences may exist at all scores or at restricted ranges (e.g., only among the middle scores). Analyses will therefore use B_0 coefficients centered on Static-99 scores of 0, 2, and 5. The score of 2 was selected because it is the score for the "typical" sex offender, defined as the estimated median score for the population of all adjudicated Canadian sex offenders (Hanson, Lloyd, Helmus, & Thornton, 2008). A score of 5 was chosen because it is in the higher range of Static-99 scores while still representing a common (i.e., high frequency) score (for percentile information, see Hanson et al., 2008). For clarity in terminology, $B_{0(0)}$, $B_{0(2)}$, and $B_{0(5)}$ were used to denote intercept coefficients centered on Static-99 scores of 0, 2, and 5, respectively.

Findings across studies were aggregated primarily using fixed effect meta-analysis (Hedges, 1994) weighted by the inverse of the variance (Hedges & Olkin, 1985). To test the variability of effects across studies, Cochran's Q statistic will be used (Hedges & Olkin, 1985). A significant Q statistic indicates that there is more variability across studies than would be expected by chance. The Q statistic is distributed as a chi-square with $k - 1$ degrees of freedom (k = the number of studies). A finding will be considered an outlier if it is the single extreme value and accounts for more than 50% of the total variance (Q). Outliers were removed from the analyses.

Given that significant variability was found across studies in the base rate estimates (B_0), moderator variables were examined to see if they account for the variability. Moderator analyses were conducted using both fixed effect meta-analysis of B_0s centered on a Static-99 of 2, and Cox regression analyses. The use of multiple analyses was hoped to minimize significant findings for variables with trivial effects, as is common with such large sample sizes.

Meta-analysis of logistic regression coefficients provided a nested approach to aggregating data, which is preferable when samples vary (Hanson & Thornton, 2000). Additionally, the *Q* statistic allows variability across studies to be measured and compared across levels of moderator variables. Weaknesses of this method include its inability to assess case-level variables such as age and year at release, and that fixed follow-up periods are required, which reduces the available sample size. For categorical moderator variables, the overall *Q* statistic was partitioned into variability across samples that could be explained by the moderator (between-level variability, which will be referred to as between-level *Q*), and unexplained variability across samples (within-level variability).

Cox regression analysis (Allison, 1984) was used to provide a different perspective and to test interactions with Static-99 scores. Cox regression estimates relative risk ratios (hazard rates) associated with one or more predictor variables from survival data with unequal follow-up times. Cox regression permits analysis of all cases (whereas logistic regression requires fixed follow-up periods), resulting in a larger sample size. The primary disadvantage of Cox regression, however, is that it assumes the shape of the survival curve is proportional across all studies. Using sample as a strata variable allows separate baseline hazard functions (i.e., recidivism rates) for each value of the stratified variable, effectively removing from the analysis the base rate variability across samples. This correction only works, however, for case-level moderator variables (e.g., age, offender type, year of release). Furthermore, to the extent that samples systematically differ on case-level moderators, using sample as a strata variable removes variability across samples that may be due to the moderator, thereby providing a more conservative test. If the moderator is highly correlated with different samples, this test may be too conservative.

Although complicated modeling of numerous variables may improve predictive accuracy, finding a parsimonious explanation of base rate variability was viewed as desirable, particularly because one of the often-cited advantages of Static-99 is its ease of use (Helmus, 2007). The principle guiding the selection of moderator variables was that the simplest solution available (i.e., using the fewest moderators) without meaningfully reducing predictive accuracy was preferable.

Rate Ratios and Odds Ratios. One measure of effect size commonly used in the current study is the exponent produced from Cox regression and logistic regression, reported as Exp(B). In Cox regression, this exponent is a rate ratio, defined as the probability of an event (recidivism) in one group divided by the probability in another group. In logistic regression, Exp(B) refers to an odds ratio, which is slightly different. An odds is the probability of an event (recidivism) divided by the probability of non-occurrence of the event. An odds ratio is the odds of an event in one group divided by the odds of that event in another group. When recidivism base rates are low (as is the case for most analyses in the current study), the odds ratio is similar to a rate ratio.

New recidivism estimates for Static-99. New recidivism norms for Static-99 were produced using random effect meta-analysis. Whereas the results of fixed effect meta-analysis are conceptually restricted to the particular set of studies in the meta-analysis, random effect meta-analysis estimates effects for the population to which the current sample of studies is a part (Hedges & Vevea, 1998). More specifically, random effect meta-analysis incorporates variability across samples into the error term, whereas fixed effect meta-analysis separates that variability. Consequently, the random effect model permits calculation of a "best estimate" for the recidivism base rate, and the variability across samples is incorporated in the confidence interval for that estimate. When variability across studies is low (Q < degrees of freedom), random effect and fixed effect meta-analysis produce identical results.

Results

Are New Recidivism Estimates Needed?

Recidivism estimates in the current samples were compared to the original estimates using survival analysis and Cox regression. Tables 9 and 10 present the sexual and violent recidivism estimates from survival analysis at 5, 10, and 15 years for the original and new datasets. Figures 2 through 7 provide a graphical display of the difference between the original and new datasets. Survival analysis was used because that is how the original recidivism estimates were calculated and there was insufficient information to calculate the fixed follow-up periods needed for logistic regression in one of the three original Static-99 samples. Recidivism estimates for the

Table 9

Sexual Recidivism Estimates from Survival Analysis

	Development Samples				Current Samples			
Score	Initial *n*	5 years	10 years	15 years	Initial *n*	5 years	10 years	15 years
0	107 (10%)	5.0	10.8	13.4	1,075 (12%)	3.4	4.5	5.5
1	150 (14%)	6.2	7.2	7.2	1,375 (16%)	4.6	6.8	10.4
2	204 (19%)	8.5	13.0	16.3	1,461 (17%)	6.0	8.6	8.6
3	206 (19%)	11.7	14.4	19.2	1,386 (16%)	8.1	14.2	18.6
4	190 (18%)	26.3	31.0	36.2	1,196 (14%)	11.8	17.6	20.4
5	100 (9%)	33.7	37.6	40.4	842 (10%)	17.9	25.2	29.0
6+	129 (12%)	38.8	44.8	52.1	1,391 (16%)	25.7	33.7	36.9
	1,086 (100%)	17.4	21.5	25.4	8,726 (100%)	10.8	15.3	18.0

Note. Mean Static-99 score was 3.0 in the development samples and 3.1 in the current samples

Table 10

Violent Recidivism Estimates from Survival Analysis

	Development Samples				Current Samples			
Score	Initial *n*	5 years	10 years	15 years	Initial *n*	5 years	10 years	15 years
0	107 (10%)	5.9	11.8	15.7	874 (12%)	6.0	9.8	16.0
1	150 (14%)	11.0	16.6	17.7	1,164 (16%)	8.8	14.6	18.8
2	204 (19%)	16.9	25.2	29.8	1,260 (17%)	13.3	18.2	21.8
3	206 (19%)	22.4	27.2	34.2	1,184 (16%)	19.6	32.4	38.4
4	190 (18%)	36.1	43.9	51.8	1,034 (14%)	26.8	37.6	43.0
5	100 (9%)	41.9	48.4	52.4	742 (10%)	35.0	46.7	48.8
6+	129 (12%)	44.4	51.4	58.7	1,202 (16%)	39.0	53.0	53.3
	1,086 (100%)	24.9	31.5	37.0	7,460 (100%)	20.8	29.8	33.9

Note. Mean Static-99 score was 3.0 in the development samples and 3.1 in the current samples

Figure 2. Comparing 5-year sexual recidivism estimates

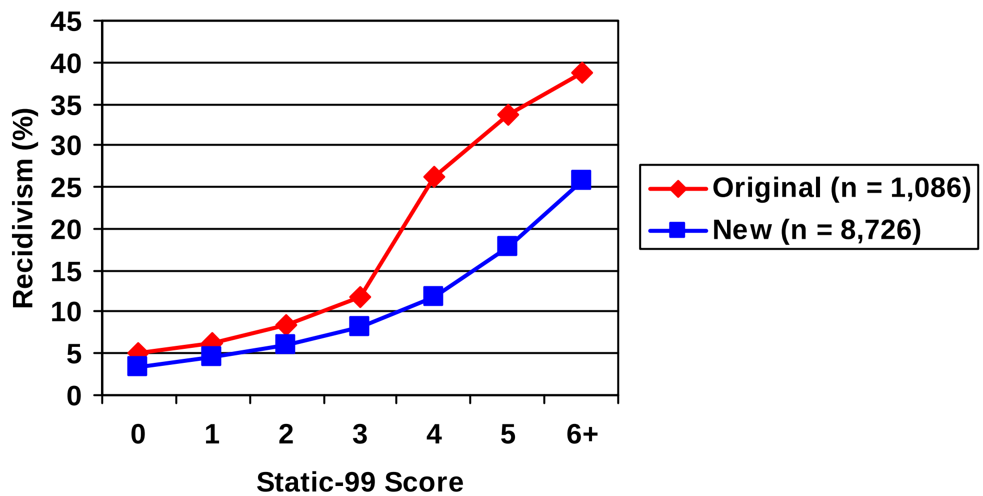


Figure 3. Comparing 10-year sexual recidivism estimates

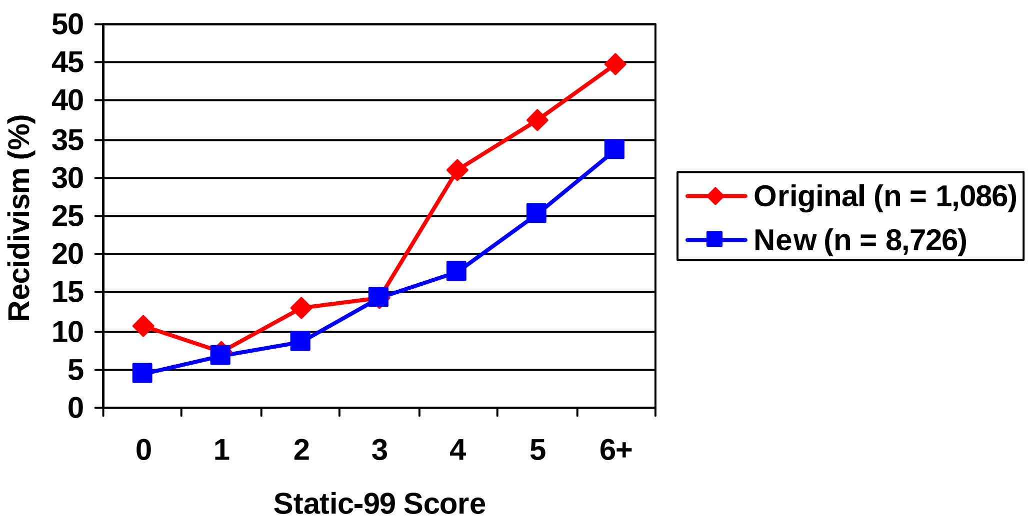


Figure 4. Comparing 15-year sexual recidivism estimates

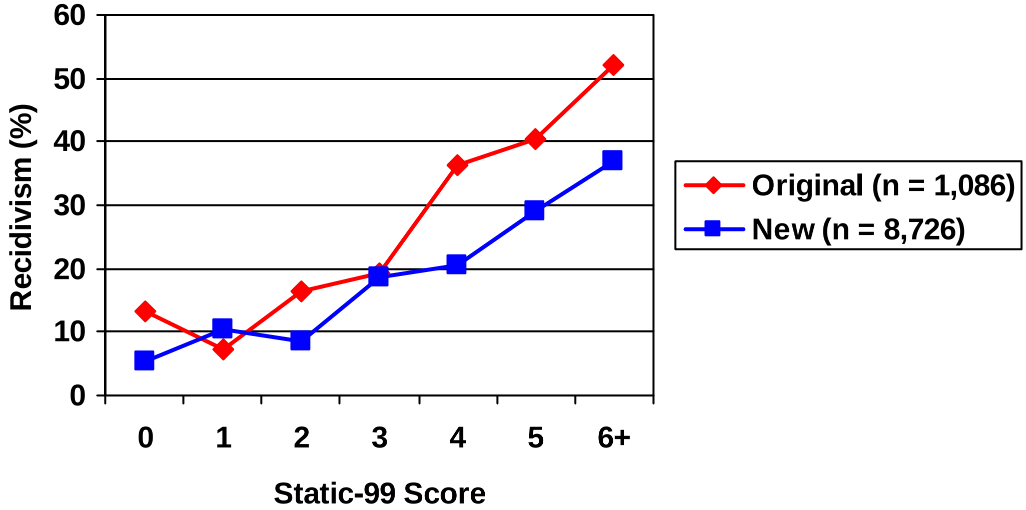


Figure 5. Comparing 5-year violent recidivism estimates

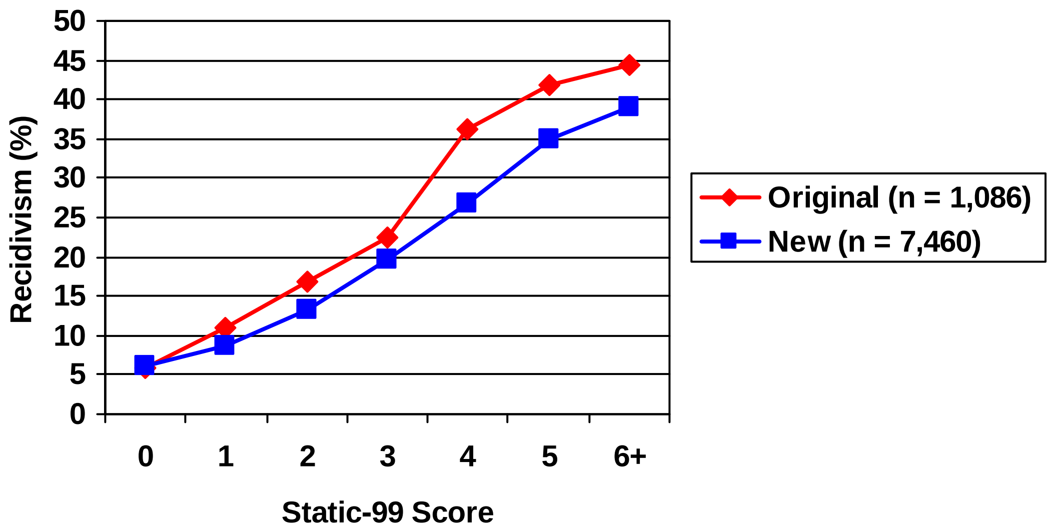


Figure 6. Comparing 10-year violent recidivism estimates

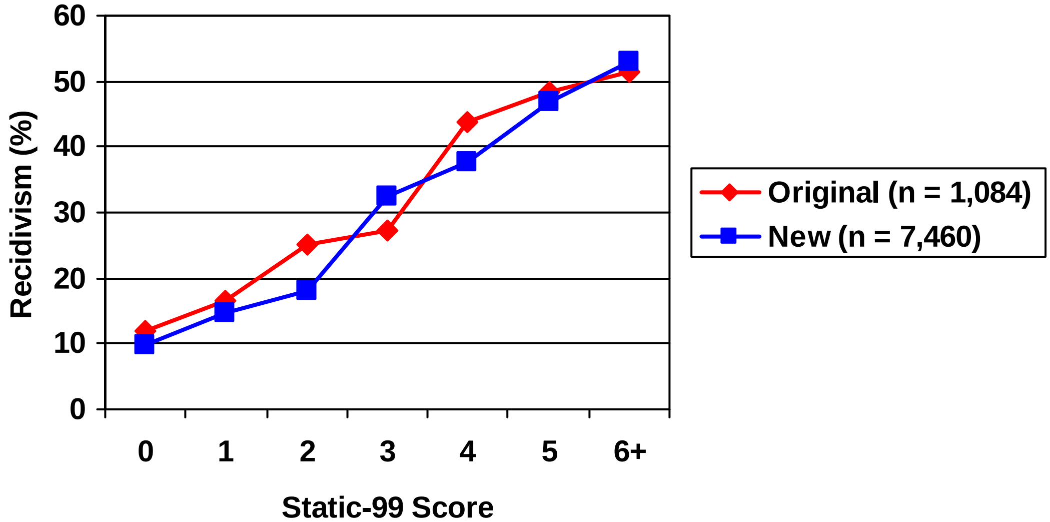
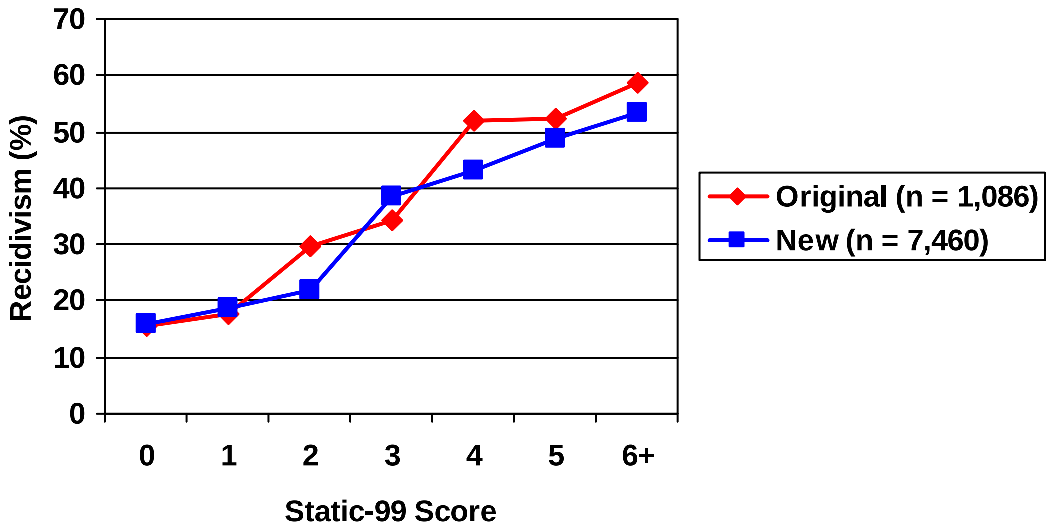


Figure 7. Comparing 15-year violent recidivism estimates



high-risk group (Static-99 scores of 6+) were collapsed into one category because that is how the original recidivism norms were presented. For sexual recidivism at 5 and 10 years, the recidivism estimate for each Static-99 score was lower in the current samples than in the original development samples. At 15 years, the sexual recidivism rate for a score of 1 was higher in the current samples than in the original samples, but all other estimates were lower in the current samples. Note that in the original estimates, the 15 year rate for a score of 1 was unusually low relative to the other scores (and was lower than the rate for a score of 0).

Examining the graphs (Figures 2 through 4), the newer samples showed noticeably lower sexual recidivism estimates compared to the original samples, and the difference is particularly striking for scores of 4 and higher. At 10 and 15 years the sexual recidivism results fluctuate in the lower Static-99 scores (particularly 1 & 3), but this appears to be due to random fluctuation in the original estimates as opposed to the current samples, which displayed more orderly patterns of increasing risk with increased Static-99 scores.

For violent recidivism at 5 years, the estimates were lower in the current samples than in the original samples for six out of seven scores. At 10 years, the estimates were lower in the current samples for five out of seven scores. At 15 years, the recidivism estimates were lower in the current samples for four out of seven comparisons. Examining the accompanying graphs (Figures 5 through 7), the rates from the current samples appear noticeably lower than the original samples at 5 years, although the magnitude of the difference is less than that found for sexual recidivism.

Similar to the sexual recidivism findings, the difference is greater for scores of 4 and higher. At 10 and 15 years, no clear pattern is evident.

Cox regression was used to test the significance of the observed differences between the current and original samples. For sexual recidivism, three samples were missing information for this analysis (Craig et al., 2006; Endrass et al., 2009; G. T. Harris et al., 2003). After controlling for Static-99, the sexual recidivism rate was significantly lower in the new samples ($k = 26$, $n = 8{,}726$) compared to the original samples ($k = 3$, $n = 1{,}084$; χ^2 change = 51.2, $df = 1$, $p < .001$), with the newer samples showing approximately 60% the recidivism rate of the original samples (Exp(B) = .59, 95% C. I. of .51 to .67). After controlling for Static-99, the violent recidivism rate was significantly lower in the new samples ($k = 21$, $n = 7{,}460$) compared to the original samples ($k = 3$, $n = 1{,}086$; χ^2 change = 15.1, $df = 1$, $p < .001$), with the newer samples showing approximately 80% the violent recidivism rate of the original samples (Exp(B) = .79, 95% C. I. of .71 to .89).

The analyses and graphs presented above indicated that sexual and violent recidivism were significantly lower in the newer samples compared to the original samples, but the magnitude of the difference is larger for sexual recidivism than for violent recidivism. One possible hypothesis for the disparity in findings between sexual and violent recidivism is that the current samples have a larger proportion of rapists than the original samples. In the current samples with offender type information ($k = 15$, $n = 6{,}335$), 37% of offenders were rapists, and 53% were child molesters (the remaining proportion consisted of non-contact and mixed offenders). The original samples ($k = 3$, $n = 938$ with offender type information), contained more child molesters (71%) and fewer rapists (29%). Although rapists and child molesters have similar rates of sexual recidivism (Hanson & Harris, 2004; Hanson et al., in press), rapists have much higher rates of violent recidivism (Hanson et al., in press). A larger proportion of rapists could, therefore, increase the violent recidivism rates without affecting the sexual recidivism rates. This means that an increased proportion of rapists could be suppressing a larger disparity in violent recidivism between the old and new samples.

A second set of Cox regression analyses were completed to compare the original and current samples, this time controlling for offender type (rapist versus child molester) in addition to Static-99 scores (in the new samples, $k = 11$, $n = 4{,}772$; from the original samples, $k = 3$, $n = 938$). Offender type significantly predicted violent recidivism, with rapists showing approximately 1.5 times the violent recidivism rate of child molesters (Exp(B) = 1.55, 95% C.I. of 1.38 to 1.73). After controlling for offender type and Static-99, the violent recidivism rate in the newer samples was still significantly lower than the original samples (χ^2 change = 19.3, df = 1, $p < .001$), and the magnitude of the difference increased, with the new samples showing approximately 74% the violent recidivism rate of the original samples (Exp(B) = .74, 95% C.I. of .64 to .84), compared to 80% in the analyses that did not control for offender type.

For sexual recidivism (in the new samples, $k = 14$, $n = 5{,}431$; in the original samples, $k = 3$, $n = 938$), when offender type, Static-99, and old versus new sample were entered in Cox regression, the difference between rapists and child molesters was non-significant (Exp(B) = .87, 95% C.I. of .75 to 1.01) and controlling for offender type in addition to Static-99 scores had a smaller effect on the difference between old and new samples, with new samples showing 56% of the recidivism rate of the original samples (Exp(B) = .56, 95% C.I. of .48 to .66).

Roughly, the sexual recidivism rate in the new samples was approximately 60% of the rate observed in the original samples. After controlling for the masking effects of offender type, the current samples displayed about three quarters of the violent recidivism rate from the original samples. These analyses indicated that the original recidivism estimates from Static-99 are overestimating recidivism in larger, more current samples, and new Static-99 recidivism norms are therefore warranted. Further analyses of the new samples will be conducted to generate new recidivism norms (the original Static-99 samples will no longer be used in these analyses).

Summary of the Recidivism Rates from the Current Samples

Logistic regression was used to summarize absolute and relative risk properties of Static-99 for each sample. Tables 11 and 12 present the logistic regression results per sample at 5 and 10 year fixed follow-up periods, with B_0s centered on Static-99 scores of 0, 2, and 5. The base rate estimates (B_0s) and their confidence intervals were converted to percentages (to view the results as log odds, see Appendixes B and C). Figures 8 through 10 provide a graphical depiction of the variability across samples in the $B_{0(0)}$s, $B_{0(2)}$s, and the $B_{0(5)}$s at 5 years. Logistic regression coefficients across samples were aggregated using fixed effect meta-analysis, displayed in Table 13.

Table 11

Static-99 Logistic Regression Analyses at Five Years with B_0s Converted to Percentages

	Static-99 *M (SD)*	*N*	% recid	B_1	B_1SE	$B_{0(0)}$	$B_{0(0)}$ 95% C.I.		$B_{0(2)}$	$B_{0(2)}$ 95% C.I.		$B_{0(5)}$	$B_{0(5)}$ 95% C.I.	
Allan et al. (2007)	2.3 (2.0)	299	11.7	0.411	0.086	3.8	2.0	7.4	8.3	5.5	12.5	23.8	16.6	32.9
Bartosh et al. (2007)	3.2 (2.2)	90	13.3	0.138	0.135	8.7	2.9	23.2	11.2	5.5	21.2	16.0	8.6	27.7
Bengtson (2008)	3.8 (2.1)	310	19.7	0.210	0.070	9.5	5.2	16.9	13.8	9.5	19.7	23.1	18.2	28.9
Bigras (2007)	2.5 (1.9)	207	9.2	0.362	0.117	3.2	1.2	8.0	6.4	3.6	11.2	16.8	10.2	26.4
Boer (2003)	3.3 (2.3)	299	3.7	0.515	0.146	0.4	0.1	2.1	1.0	0.3	3.4	4.6	2.4	8.5
Bonta & Yessine (2005)	5.2 (2.0)	101	18.8	0.265	0.139	5.1	1.0	22.1	8.4	2.8	22.5	16.8	10.4	26.0
Brouillette-Alarie & Proulx (2008)	3.8 (2.2)	199	14.6	0.360	0.095	3.4	1.3	8.8	6.7	3.5	12.5	17.5	12.3	24.3
Cortoni & Nunes (2007)	3.2 (1.9)	17	.0	-	-	1.8	0.03	43.6	1.8	0.03	43.6	1.8	0.03	43.6
Craissati et al. (2008)	2.3 (2.0)	200	7.5	0.358	0.114	2.7	1.0	6.9	5.4	2.8	9.9	14.2	8.2	23.6
de Vogel et al. (2004)	6.1 (1.7)	100	26.0	0.459	0.150	1.8	0.2	12.3	4.4	1.1	16.6	15.6	8.6	26.6
Eher et al. (2008)	2.3 (1.7)	151	2.0	0.695	0.255	0.2	0.01	2.7	0.6	0.1	4.6	4.8	1.4	14.9
Endrass et al. (2009)	3.5 (1.7)	95	8.4	0.470	0.205	1.3	0.2	9.2	3.4	0.9	11.3	12.4	6.0	23.4
Epperson (2003)	2.8 (2.2)	151	10.6	0.384	0.116	3.0	1.0	8.4	6.2	3.1	12.2	17.4	10.6	27.2
Haag (2005)	3.9 (2.0)	198	19.7	0.296	0.094	6.4	2.7	14.8	11.1	6.4	18.5	23.2	17.3	30.4
Hanson et al. (2007)	3.1 (2.1)	31	.0	-	-	0.8	0.02	29.4	0.8	0.02	29.4	0.8	0.02	29.4

Table continues.

Table 11 continued.

	Static-99 *M* (*SD*)	*N*	% recid	B_1	B_1SE	$B_{0(0)}$	$B_{0(0)}$ 95% C.I.		$B_{0(2)}$	$B_{0(2)}$ 95% C.I.		$B_{0(5)}$	$B_{0(5)}$ 95% C.I.	
Harkins & Beech (2007)	2.8 (2.2)	198	9.6	0.377	0.108	2.7	1.0	7.3	5.6	2.9	10.6	15.6	9.9	23.6
Hill et al. (2008)	5.0 (1.8)	73	11.0	0.172	0.212	4.8	0.5	35.2	6.6	1.4	26.1	10.6	5.2	20.2
Johansen (2007)	3.0 (2.0)	272	5.9	0.234	0.120	2.7	1.0	7.4	4.3	2.2	8.1	8.3	4.8	13.9
Knight & Thornton (2007)	4.5 (2.2)	433	24.7	0.241	0.054	9.4	5.5	15.6	14.4	10.2	19.9	25.7	21.6	30.2
Långström (2004)	2.4 (2.0)	1278	5.4	0.328	0.054	2.0	1.3	3.2	3.9	2.9	5.2	9.7	7.5	12.5
Langton (2003)	3.3 (2.1)	226	10.2	0.250	0.106	4.3	1.7	10.5	6.8	3.8	12.1	13.4	8.8	19.9
Milton (2003)	4.8 (2.0)	93	16.1	0.428	0.160	1.9	0.3	11.7	4.4	1.2	14.8	14.3	8.1	23.8
Nicholaichuk (2001)	4.5 (2.0)	168	22.6	0.391	0.109	4.1	1.3	12.2	8.5	4.0	17.1	23.0	16.9	30.5
Saum (2007)	2.0 (1.6)	175	31.4	0.561	0.120	11.6	6.4	20.1	28.7	22.0	36.5	68.4	51.5	81.5
Swinburne Romine et al. (2008)	1.9 (1.8)	570	8.4	0.279	0.073	4.6	2.9	7.4	7.8	5.8	10.4	16.4	11.2	23.5
Ternowski (2004)	2.1 (1.9)	247	6.5	0.246	0.114	3.5	1.5	7.9	5.7	3.3	9.5	11.2	6.0	19.7
Wilson et al. (2007a & b)	5.5 (2.1)	100	12.0	0.027	0.148	10.5	2.0	39.7	11.0	3.6	29.1	11.8	6.7	20.1

Note. In samples with no recidivists, B_0 was estimated but B_1 could not be computed.

Table 12

Static-99 Logistic Regression Analyses at Ten Years with B_0s Converted to Percentages

	Static-99 *M (SD)*	*N*	% recid	B_1	B_1SE	$B_{0(0)}$	$B_{0(0)}$ 95% C. I.		$B_{0(2)}$	$B_{0(2)}$ 95% C.I.		$B_{0(5)}$	$B_{0(5)}$ 95% C.I.	
Allan et al. (2007)	2.1 (1.7)	25	20.0	0.780	0.387	3.2	0.3	29.3	13.4	3.8	37.6	61.7	19.7	91.4
Bengtson (2008)	3.8 (2.1)	291	28.5	0.244	0.067	13.2	7.7	21.7	19.8	14.4	26.6	33.9	27.9	40.5
Boer (2003)	3.2 (2.3)	295	7.8	0.408	0.100	1.5	0.5	4.4	3.4	1.7	6.8	10.7	7.1	15.7
Brouillette-Alarie & Proulx (2008)	3.8 (2.2)	110	20.9	0.368	0.115	5.3	1.7	14.9	10.4	5.1	19.9	25.9	17.5	36.4
Craig et al. (2006)	2.4 (1.7)	66	13.6	0.148	0.201	9.8	2.9	28.0	12.7	6.4	23.7	18.5	6.7	41.7
Craissati et al. (2008)	1.7 (1.7)	66	9.1	0.190	0.216	6.4	1.9	19.6	9.1	4.1	19.0	15.1	4.0	42.9
de Vogel et al. (2004)	6.3 (1.8)	71	38.0	0.509	0.170	2.2	0.2	18.0	6.0	1.2	24.2	22.5	12.2	37.9
Epperson (2003)	3.6 (2.5)	36	22.2	1.068	0.380	0.2	.001	10.4	1.2	0.1	19.3	23.5	8.2	51.3
Harkins & Beech (2007)	2.9 (2.2)	129	16.3	0.392	0.114	4.6	1.7	12.1	9.6	5.1	17.4	25.7	16.8	37.3
Hill et al. (2008)	5.1 (1.8)	54	18.5	0.050	0.201	14.9	2.0	60.0	16.3	4.5	44.6	18.4	10.2	31.1
Johansen (2007)	3.5 (2.2)	62	12.9	-0.004	0.177	13.1	3.5	38.4	13.0	5.7	27.0	12.8	5.6	26.7
Knight & Thornton (2007)	4.4 (2.3)	353	30.0	0.206	0.054	14.3	8.7	22.6	20.1	14.8	26.9	31.9	27.0	37.2
Långström (2004)	2.5 (2.0)	353	7.4	0.484	0.091	1.5	0.6	3.5	3.8	2.1	6.6	14.4	9.7	20.7
Langton (2003)	3.1 (2.1)	47	12.8	0.418	0.225	3.0	0.4	20.4	6.6	1.8	21.9	19.8	8.8	38.8
Milton (2003)	5.0 (2.0)	68	25.0	0.404	0.161	3.6	0.6	19.9	7.7	2.2	23.7	22.0	13.1	34.5
Nicholaichuk (2001)	4.7 (2.1)	59	25.4	0.274	0.160	8.1	1.5	33.3	13.2	4.5	32.9	25.8	15.9	38.9
Swinburne Romine et al. (2008)	1.9 (1.8)	543	11.2	0.232	0.069	7.1	4.8	10.5	10.9	8.5	13.8	19.6	13.6	27.4
Wilson et al. (2007a & b)	5.6 (2.5)	15	6.7	17.404	3473.3	<.01	<.01	>99.9	<.01	<.01	>99.9	<.01	<.01	>99.9

Figure 8. Estimated 5-year recidivism rate for a Static-99 Score of 0 ($B_{0(0)}$) for each sample and combined weighted average

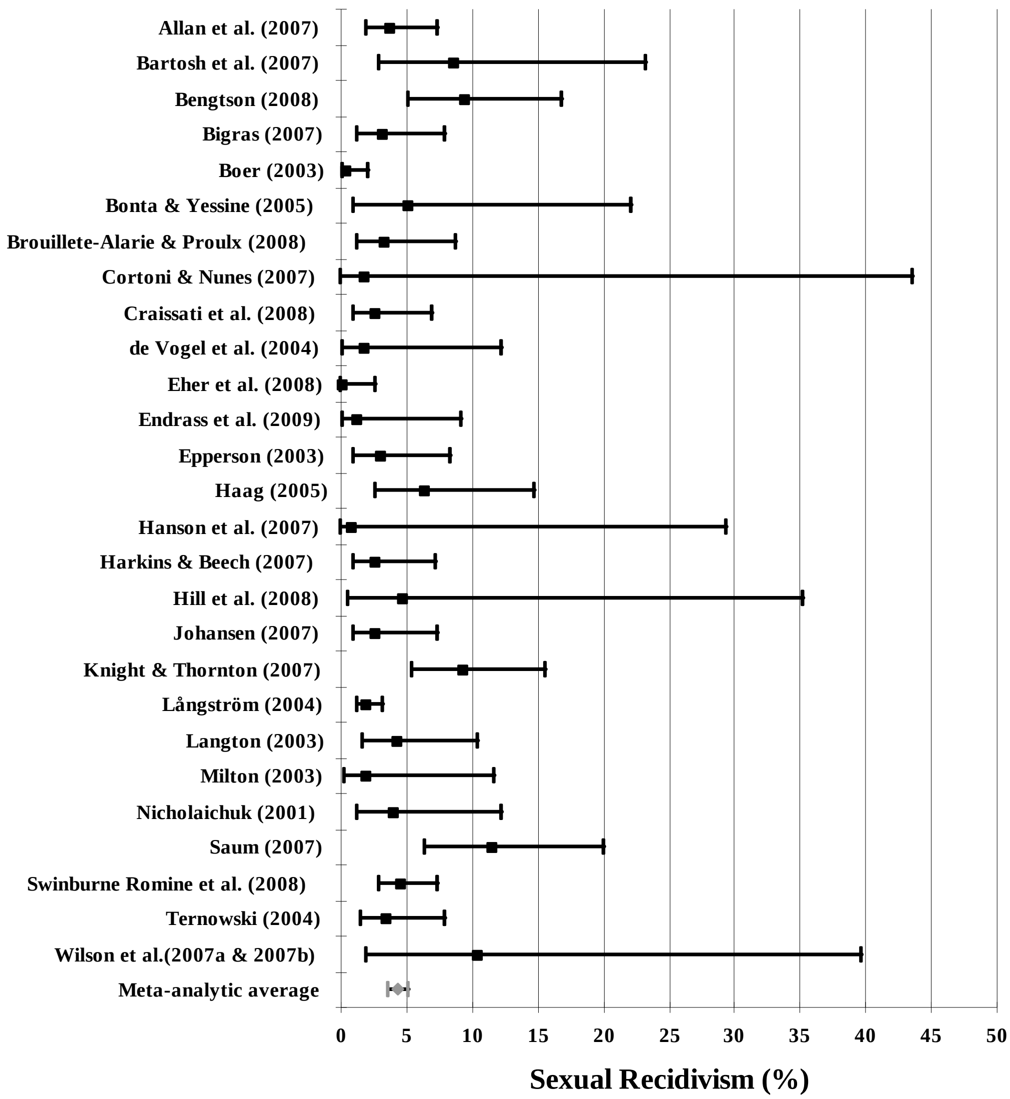


Figure 9. Estimated 5-year recidivism rate for a Static-99 Score of 2 ($B_{0(2)}$) for each sample and combined weighted average

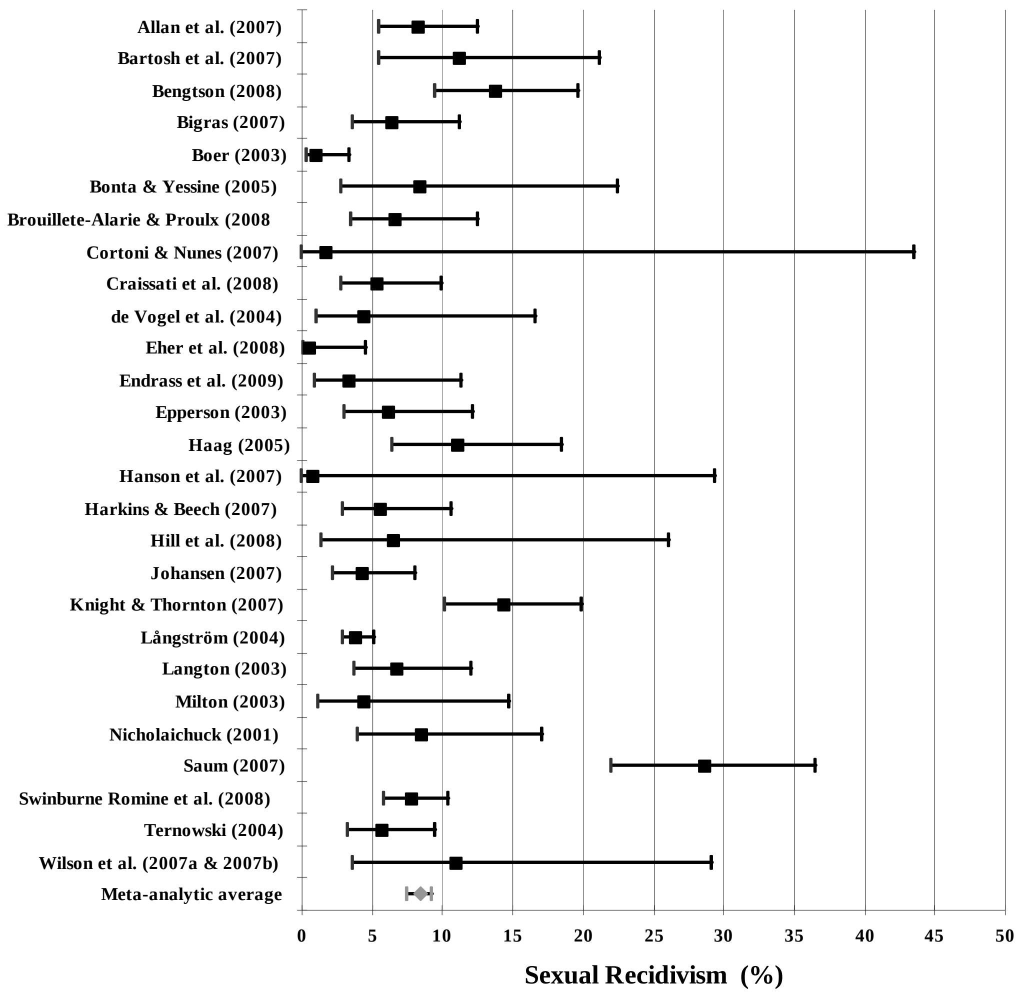


Figure 10. Estimated 5-year recidivism rate for a Static-99 Score of 5 ($B_{0(5)}$) for each sample and combined weighted average

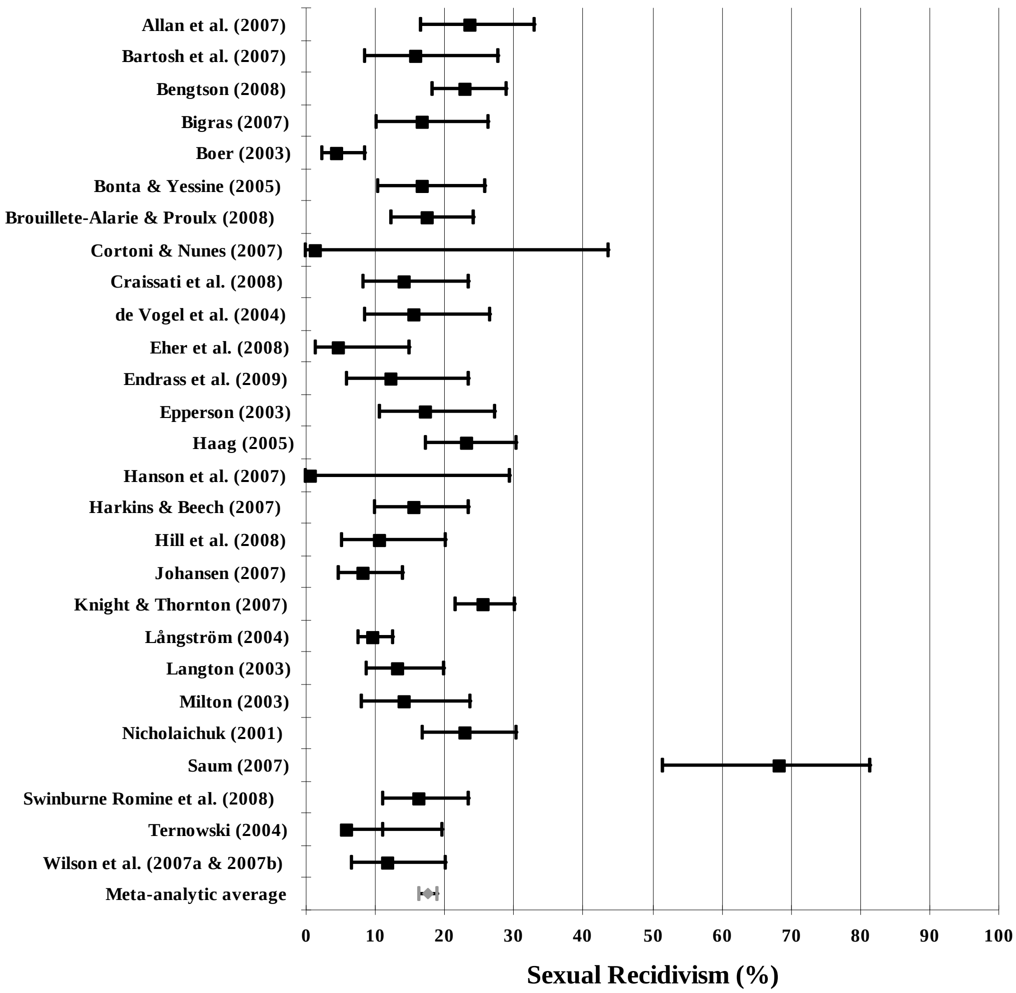


Table 13

Fixed Effect Meta-Analysis of Logistic Regression Coefficients for Sexual Recidivism

	M	95 % CI	*Q*	*k*	*n*
Five Years					
B_1	0.312	0.272 – 0.351	24.91	25	6,233
B_0					
Centered 0	4.3%	3.6 – 5.1	59.73***	27	6,281
Centered 2	8.4%	7.5 – 9.3	149.09***	27	6,281
Centered 5	17.6%	16.3 – 19.0	144.71***	27	6,281
Ten Years					
B_1	0.290	0.237 – 0.342	23.92	17	2,628
B_0					
Centered 0	7.1%	5.7 – 8.8	42.74***	17	2,628
Centered 2	11.8%	10.3 – 13.4	56.10***	17	2,628
Centered 5	24.6%	22.4 – 27.0	57.73***	17	2,628

Note. Total sample size and number of studies differ between B_0 and B_1 because B_1 cannot be computed without recidivists.

At five years, the weighted average B_1 was .31 (95% C.I. of .27 to .35). Converting the log odds ratio to an odds ratio, that means that for each one-point increase in Static-99 score, the odds of recidivism increases by 1.37. The *Q* was non-significant ($Q = 24.91$, $p > .25$), indicating that the variability in predictive accuracy across studies was no more than would be expected by chance.

The B_0s, however, showed significant variability across samples regardless of where they were centered on Static-99. Among the individual studies, the predicted recidivism rate for a Static-99 score of 0 varied between 0.2% and 11.6%, with a weighted average of 4.3% (95% C.I. of 3.6 to 5.1). The predicted recidivism rate for a Static-99 score of 2 varied between 0.6% and 28.7%, with a weighted average of

8.4% (95% C.I. of 7.5 to 9.3). The predicted recidivism rate for a Static-99 score of 5 varied between 0.8% and 68.4%, with a weighted average of 17.6% (95% C.I. of 16.3 to 19.0). For $B_{0(0)}$, $B_{0(2)}$, and $B_{0(5)}$, the highest base rate was from Saum (2007). For $B_{0(0)}$ and $B_{0(2)}$, the lowest base rate was from Eher and colleagues (2009), whereas for $B_{0(5)}$, the lowest base rate was from Hanson and colleagues (2007).

Fewer samples had follow-up information at 10 years (for B_0s, k = 18, versus k = 27 for 5 year data) and most of the samples had substantially fewer cases (for B_0s, n = 2,643, versus n = 6,281 at 5 years). Two samples had unusual findings for the predictive accuracy of Static-99, although both had small sample sizes. Static-99 did not predict recidivism in Johansen's (2007) sample (n = 62), and the B_1 was slightly negative, resulting a small decrease in the predicted recidivism rates for higher Static-99 scores (a predicted recidivism rate of 13.1% for a score of 0, and a 12.8% for a score of 5).

Wilson and colleagues' (2007a & b) sample had only 15 cases with a 10 year follow-up, with one sexual recidivist (with a Static-99 score of 9). The small sample size and high risk recidivist led to an artificially inflated B_1 of 17.40 (as a comparison, all other B_1s ranged between -0.004 and 1.07). Additionally, the predicted recidivism rates for Static-99 scores of 0, 2, and 5 in this sample were all less than one one-hundredth of a percent, with incredibly large confidence intervals (ranging from less than one one-thousandth of a percent to over 99.99 percent). Because the small sample sizes and large standard errors rendered these results both illogical and uninterpretable, this sample was excluded from the meta-analysis at 10 years.

The average B_1 from the 10-year meta-analysis was .29 (95% C.I. of .24 to .34). Converting the log odds ratio to an odds ratio, that means that for each one-point increase in Static-99 score, the odds of recidivism increases by 1.34. The *Q* was non-significant ($Q = 23.92$, $p > .10$), indicating that the variability in predictive accuracy across studies was no more than would be expected by chance.

Similar to the five year findings, the B_0s showed significant variability across samples regardless of where they were centered. Among the individual studies, the predicted recidivism rate for a Static-99 score of 0 varied between 0.2% and 14.9%, with a weighted average of 7.1% (95% C.I. of 5.7 to 8.8). The predicted recidivism rate for a Static-99 score of 2 varied between 1.2% and 20.1%, with a weighted average of 11.8% (95% C.I. of 10.3 to 13.4). The predicted recidivism rate for a Static-99 score of 5 varied between 10.7% and 61.7%, with a weighted average of 24.6% (95% C.I. of 22.4 to 27.0). The highest base rates were from Hill and colleagues (2008) for $B_{0(0)}$, Knight and Thornton (2007) for $B_{0(2)}$, and Allan and colleagues (2007) for $B_{0(5)}$. Note that Saum (2007), which consistently had the highest base rate at 5 years, had no 10 year information and was therefore excluded from these analyses. Similarly, at 5 years the lowest base rates were from Eher and colleagues (2009) and Hanson et al. (2007), which were both not available for analysis at 10 years. For $B_{0(0)}$ and $B_{0(2)}$, the lowest base rate was from Epperson (2003), whereas for $B_{0(5)}$, the lowest base rate was from Boer (2003).

Although the variability across studies (measured by *Q*) was significant for the B_0s at all scores examined, there was more variability in predicted recidivism rates for Static-99 scores of 2 and 5 than there was for scores of 0. Because analyses centered on the median Static-99 score of 2 appeared to effectively capture the variability across samples, subsequent analyses of B_0s were typically centered only on this score. The significant variability in B_0s demonstrated that the assumption of stability across samples in the recidivism estimates for Static-99 is untenable. These findings warranted examination of some possible moderator variables.

Before the moderator analyses were conducted, Saum's (2007) sample was removed because it was considered an outlier in the 5 year analyses (no cases were available for the 10 year analyses). In addition to having the highest predicted recidivism rate for Static-99 scores of 0, 2, and 5, the recidivism estimates (particularly for scores of 2 and 5) are sufficiently distanced from those found in other samples that they do not appear plausible (see Figures 8 through 10). For Static-99 scores of 5, the confidence intervals from Saum's sample do not overlap with the confidence intervals from any other sample. In the meta-analysis of $B_{0(2)}$, Saum's study represented 49.7% of the overall Q (74.12 of 149.09), which suggested it was a statistical outlier.

A possible explanation for the unusually high recidivism rates in Saum's (2007) study is disproportionate data retention. Offenders in this sample were treated by the North Dakota Department of Human Services from 1988 to 1998. Based on the completion date of the thesis, data was likely collected around 2004-2006, meaning that all files were minimally 6-16 years old at the time of data collection. Saum noted that some treatment programs lost or destroyed some of their records in the process of changing locations. It is plausible that the purged records would disproportionately include files that had been inactive for some time (i.e., non-recidivists). The biasing effects of disproportionate data purging in Canadian CPIC records have been found to artificially increase base rates when records are retrospectively sampled (Hanson & Nicholaichuk, 2000), which may have occurred in Saum's study as well.

Moderator Analyses

Numerous moderator variables were analyzed using Cox regression and fixed effect meta-analysis of logistic regression coefficients at fixed five year follow-ups. Table 14 presents the Cox regression analyses for all moderators (with the exception of age at release, which is presented in Table 16) after controlling for Static-99 scores. Additionally, interactions between each dichotomous or linear moderator and Static-99 scores were also tested. Results are presented for the interaction term only (not the original moderator variable after controlling for the interaction term).

Table 15 presents the meta-analytic findings for all moderators except age and year of release. For each moderator, fixed effect meta-analytic results are presented overall and for each level of the moderator. When outliers were identified, results were presented with and without the outlying study (with the exception of Saum, 2007, which was excluded from all analyses). Analyses of B_0 are centered on a Static-99 score of 2. The variability (Q) within each level of the moderator was subtracted from the overall Q statistic, and the remaining Q value represents the variability across studies due to the moderator variable (between-level Q; tested on a chi-square distribution with the number of levels minus one as the degrees of freedom). For each analysis, there is a separate table in Appendix D that presents the findings from each individual sample or subsample used for each level of the moderator. Each moderator will be discussed in turn, with ancillary analyses presented as needed.

Table 14

Cox Regression Moderator Analyses for Sexual Recidivism After Controlling for Static-99

				χ^2 change			Regression coefficient			Rate ratio	
Variable	*N*	*N* recid	*k*	change	*df*	*p*	*b*	*SE*	*p*	Exp(B)	95% C.I.
Recidivism criteria	8,557	1,064	25	4.048	1	.044	.125	.062	.044	1.339	1.306 – 1.372
Interaction with Static-99				10.524	1	.001	-.081	.025	.001	.922	.878 – .968
Number of recidivism sources	8,557	1,064	25	4.693	1	.030	.059	.027	.028	1.061	1.006 – 1.118
Interaction with Static-99				0.710	1	.400	-.009	.011	.400	.991	.970 – 1.012
More than 1 recidivism source	8,557	1,064	25	1.134	1	.252	.076	.066	.250	1.079	.948 – 1.228
Interaction with Static-99				0.889	1	.346	-.025	.027	.346	.975	.925 – 1.028
Used National records	8,557	1,064	25	0.004	1	.951	-.005	.088	.951	.995	.831 – 1.182
Interaction with Static-99				2.055	1	.152	.052	.037	.155	1.054	.980 – 1.132
Used street time	8,557	1,064	25	1.809	1	.179	.105	.077	.174	1.111	.955 – 1.292
Interaction with Static-99				1.192	1	.275	-.035	.032	.275	.966	.908 – 1.028
Cited Coding Rules	8,557	1,064	25	0.095	1	.758	-.021	.069	.757	.979	.855 – 1.120
Interaction with Static-99				0.098	1	.754	-.009	.028	.754	.991	.938 – 1.047

Table continues.

Table 14 continued.

Variable	*N*	*N* recid	*k*	χ^2 change: change	*df*	*p*	Regression coefficient: *b*	*SE*	*p*	Rate ratio: Exp(B)	95% C.I.
Provincial versus federal	3,137	356	11	0.668	1	.414	-.102	.124	.410	.903	.707 – 1.152
Interaction with Static-99				2.740	1	.098	-.082	.050	.097	.920	.834 – 1.015
Child molester versus rapist	5,370	557	13	1.862	1	.172	-.117	.086	.173	.889	.751 – 1.053
Interaction with Static-99				0.329	1	.556	.021	.036	.566	1.021	.951 – 1.095
Child molester versus rapist	5,367	557	13	0.064	1	.800	-.024	.093	.800	.977	.814 – 1.172
Interaction with Static-99 -strata				0.080	1	.777	.011	.037	.777	1.011	.940 – 1.087
Country (U.S.)	8,557	1,064	25	6.517	4	.164					
Canada							-.124	.080	.123	.883	.754 – 1.034
UK							-.139	.133	.296	.870	.670 – 1.130
Europe							-.172	.084	.041	.842	.714 – .993
New Zealand							.113	.159	.447	1.120	.820 – 1.529
Year of Release	8,556	1,064	25	15.908	1	<.001	-.013	.003	<.001	.987	.981 – .993
Interaction with Static-99				11.891	1	.001	.004	.001	.001	1.004	1.002 – 1.007
Year of Release - strata	8,473	1,064	25	1.340	1	.247	.009	.008	.247	1.009	.994 – 1.025
Interaction with Static-99 -strata				11.350	1	.001	.005	.001	.001	1.005	1.002 – 1.007

Table continues.

Table 14 continued.

				χ^2 change			Regression coefficient			Rate ratio	
Variable	*N*	*N* recid	*k*	change	*df*	*p*	*b*	*SE*	*p*	Exp(B)	95% C.I.
Aboriginal	2,569	357	8	3.763	1	.052	.296	.147	.045	1.344	1.007 – 1.795
Interaction with Static-99				1.149	1	.284	.082	.077	.288	1.085	.933 – 1.262
Aboriginal - strata	2,566	375	8	3.126	1	.077	.284	.157	.070	1.328	.977 – 1.085
Interaction with Static-99 -strata				0.549	1	.459	.058	.079	.461	1.060	.908 – 1.237
Non White	2,166	381	8	2.152	1	.142	.207	.138	.133	1.229	.939 – 1.610
Interaction with Static-99				1.199	1	.274	.068	.062	.274	1.070	.948 – 1.208
Non White - strata	2,163	381	8	3.216	1	.073	.285	.155	.065	1.330	.983 – 1.801
Interaction with Static-99 -strata				0.777	1	.378	.055	.063	.379	1.057	.934 – 1.195
NonAboriginal Minority	1,912	339	7	1.385	1	.239	.270	.221	.221	1.310	.850 – 2.020
Interaction with Static-99				0.001	1	.980	.002	.086	.980	1.002	.846 – 1.187
NonAboriginal Minority - strata	1,910	339	7	0.518	1	.471	.164	.223	.461	1.179	.761 – 1.826
Interaction with Static-99 -strata				<0.001	1	.986	-.002	.088	.986	.998	.841 – 1.186
Treatment status (Didn't start)	1,801	233	5	3.278	2	.194					
Completed							-.028	.178	.874	.972	.685 – 1.379
Dropped out							.248	.195	.205	1.281	.873 – 1.879
Treatment status (Didn't start) - strata	1,801	233	5	4.244	2	.120					
Completed - strata							.013	.196	.948	1.013	.689 – 1.489
Dropped out - strata							.382	.213	.074	1.464	.964 – 2.225

Table continues.

Table 14 continued.

				χ^2 change			Regression coefficient			Rate ratio	
Variable	*N*	*N* recid	*k*	change	*df*	*p*	*b*	*SE*	*p*	Exp(B)	95% C.I.
Start treatment	3,244	497	13	0.009	1	.925	.013	.141	.926	1.013	.769 – 1.375
Interaction with Static-99				1.174	1	.279	.064	.059	.281	1.066	.949 – 1.198
Start treatment - strata	3,170	497	12	0.234	1	.628	.076	.158	.630	1.079	.792 – 1.470
Interaction with Static-99 -strata				1.006	1	.316	.062	.063	.318	1.064	.942 – 1.203
Complete treatment	1,801	233	5	1.649	1	.199	-.171	.133	.198	.843	.649 – 1.094
Interaction with Static-99				1.594	1	.207	.070	.055	.208	1.072	.962 – 1.195
Complete treatment - strata	1,801	233	5	0.974	1	.324	-.161	.164	.326	.851	.617 – 1.174
Interaction with Static-99 -strata				2.110	1	.146	.082	.057	.148	1.085	.971 – 1.212
Setting (Corrections)	8,557	1,064	25	60.726	2	<.001					
Mixed							.564	.101	<.001	1.758	1.443 – 2.143
Mental Health							.575	.084	<.001	1.777	1.507 – 2.096
SampleType (Preselected HR)	6,544	585	19	85.673	3	<.001					
Routine CSC							-.585	.103	<.001	.557	.455 – .682
Community							-.131	.103	.202	.877	.717 – 1.073
Routine European prison							-.827	.106	<.001	.437	.355 – .539

Notes. Interaction values control for Static-99 and the moderator variable. For categorical variables, the value in brackets after the variable name is the reference category.

Table 15

Meta-Analysis of Categorical Moderator Variables, with B0(2) Converted to a Percentage

	$B_{0(2)}$ *M*	95% C.I.	*Q*	*n*	*k*
Recidivism criteria					
Overall	7.1	6.4 – 8.0	74.97***	6,106	26
Charges	9.3	7.8 – 10.9	25.81**	2,157	11
Convictions	5.8	5.0 – 6.7	32.65**	3,949	15
Between-level Q			**16.51*** **		
Number of recidivism sources					
Overall	7.1	6.4 – 8.0	74.97***	6,106	26
One	6.8	6.0 – 7.8	53.71***	4,579	18
Two	5.6	4.0 – 7.8	1.65	863	5
Four	11.0	8.1 – 14.8	9.52**	664	3
Between-level Q			**10.09** **		
More than one recidivism source					
Overall (excluding outlier)[a]	6.6	5.8 – 7.4	58.17***	5,673	25
One	6.8	6.0 – 7.8	53.71***	4,579	18
Two or more	5.5	4.1 – 7.3	2.62	1,094	7
Between-level Q			1.84		
Overall	7.1	6.4 – 8.0	74.97***	6,106	26
One	6.8	6.0 – 7.8	53.71***	4,579	18
Two or more	8.0	6.4 – 10.1	19.78**	1,527	8
Between-level Q			.14		
Used national criminal records					
Overall	7.1	6.4 – 8.0	74.97***	6,106	26
No	6.7	5.3 – 8.4	3.42	1,240	4
Yes	7.3	6.4 – 8.3	71.12***	4,866	22
Between-level Q			.43		

Table continues.

Table 15 continued.

	$B_{0(2)}$ *M*	95% C.I.	*Q*	*n*	*k*
Used street time?					
Overall	7.1	6.4 – 8.0	74.97***	6,106	26
No	6.6	5.8 – 7.4	57.01***	5,569	23
Yes	13.4	9.6 – 18.5	3.22	537	3
Between-level Q			14.74***		
Did they cite the coding rules?					
Overall	7.1	6.4 – 8.0	74.97***	6,106	26
No	6.4	5.0 – 8.1	2.62	1,516	10
Yes	7.4	6.5 – 8.4	71.21***	4,590	16
Between-level Q			1.14		
Canadian provincial versus federal					
Overall	6.8	5.4 – 8.5	16.67	1,781	14
Provincial	5.9	3.8 – 9.1	1.19	418	4
Federal	7.2	5.5 – 9.4	14.92	1,363	10
Between-level Q			.56		
Offender type					
Overall	7.2	6.2 – 8.3	50.13***	3,474	24
Child molester	8.0	6.7 – 9.6	30.27**	2,014	12
Rapist	5.5	4.2 – 7.3	14.81	1,460	12
Between-level Q			**5.05***		

Table continues.

Table 15 continued.

	$B_{0(2)}$ *M*	95% C.I.	*Q*	*n*	*k*
Country					
Overall (excluding outlier)[b]	6.7	5.9 – 7.5	62.17***	5,796	25
United States	8.9	7.4 – 10.8	14.61**	1,516	5
Canada	6.8	5.4 – 8.5	16.32	1,793	11
United Kingdom	5.4	3.5 – 8.1	.11	491	3
Europe	3.8	2.9 – 5.0	3.74	1,697	5
New Zealand	8.3	5.5 – 12.5	-	299	1
Between-level Q			**27.39*** **		
Overall	7.1	6.4 – 8.0	74.97***	6,106	26
United States	8.9	7.4 – 10.8	14.61**	1,516	5
Canada	6.8	5.4 – 8.5	16.32	1,793	11
United Kingdom	5.4	3.5 – 8.1	.11	491	3
Europe	5.8	4.6 – 7.2	32.79***	2,007	6
New Zealand	8.3	5.5 – 12.5	-	299	1
Between-level Q			11.14*		
Aboriginal					
Overall	7.4	5.9 – 9.2	21.49	1,588	15
Nonaboriginal	6.8	5.3 – 8.6	8.61	1,411	8
Aboriginal	14.2	7.3 – 25.9	8.68	177	7
Between-level Q			**4.20***		
NonWhite					
Overall (excluding outlier)[c]	9.4	7.7 – 11.4	18.57	1,602	13
White	8.9	7.2 – 11.0	7.80	1,390	6
NonWhite	14.5	8.0 – 24.8	8.37	212	7
Between-level Q			2.40		
Overall	8.9	7.3 – 10.8	28.42*	1,830	14
White	8.4	6.8 – 10.3	17.01*	1,618	7
NonWhite	14.5	8.0 – 24.8	8.37	212	7
Between-level Q			3.04		

Table continues.

Table 15 continued.

	$B_{0(2)}$ *M*	95% C.I.	*Q*	*n*	*k*
NonAboriginal and NonWhite					
Overall (excluding outlier)[c]	9.2	7.5 – 11.3	14.58	1,466	11
White	8.9	7.2 – 11.0	7.80	1,390	6
NonAboriginal NonWhite	17.4	7.2 – 36.3	4.62	76	5
Between-level Q			2.16		
Overall	8.4	7.1 – 10.6	24.20	1,694	12
White	8.4	6.8 – 10.3	17.01*	1,618	7
NonAboriginal NonWhite	17.4	7.2 – 36.3	4.62	76	5
Between-level Q			2.57		
Treatment					
Overall	7.9	6.5 – 9.6	14.26	1,506	12
Did not start treatment	8.8	5.9 – 13.0	2.17	312	4
Started but dropped out	9.6	6.6 – 13.7	1.08	329	3
Completed treatment	6.6	4.9 – 8.8	8.05	865	5
Between-level Q			2.96		
Started treatment?					
Overall	8.1	6.1 – 9.5	38.12**	2,609	18
No	9.4	6.4 – 13.5	3.02	360	5
Yes	7.9	6.6 – 9.4	34.46***	2,249	13
Between-level Q			.64		
Completed treatment?					
Overall	7.9	6.4 – 9.6	11.08	1,506	9
No	9.2	7.0 – 11.9	.27	641	4
Yes	6.6	4.9 – 8.8	8.05	865	5
Between-level Q			2.76		
Setting					
Overall	7.1	6.4 – 8.0	86.07***	6,106	27
Mostly corrections	6.2	5.5 – 7.1	43.86**	5,164	21
Mostly mental health	13.9	9.5 – 20.0	14.85**	559	4
Mixed	13.2	9.2 – 18.6	.93	383	2
Between-level Q			**26.43*** **		

Table continues.

Table 15 continued.

	$B_{0(2)}$ *M*	95% C.I.	*Q*	*n*	*k*
Sample Type					
Overall	7.2	6.3 – 8.2	69.52***	4,766	21
Preselected high risk	11.8	9.6 – 14.4	8.40	1,576	9
Routine CSC	5.3	3.7 – 7.7	8.79	843	5
Routine European prison	3.7	2.8 – 4.9	3.16	1,524	3
Community	7.1	5.5 – 9.2	2.94	823	4
Between-level Q			**46.23*** **		
Sample Type Revised - 1					
Overall (excluding outlier)[d]	6.7	6.0 – 7.6	60.18***	5,908	26
Routine Corrections	5.0	4.2 – 5.9	23.68*	3,191	13
Preselected high risk	10.0	6.4 – 15.5	1.55	398	3
Remaining Samples	8.4	7.1 – 9.8	14.24	2,319	10
Between-level Q			**20.71*** **		
Overall	7.1	6.4 – 8.0	86.07***	6,106	27
Routine Corrections	5.0	4.2 – 5.9	23.68*	3,191	13
Preselected high risk	15.2	10.8 – 20.9	10.19*	596	4
Remaining Samples	8.4	7.1 – 9.8	14.24	2,319	10
Between-level Q			**37.96*** **		
Sample Type Revised - 2					
Overall (excluding outlier)[d]	6.7	6.0 – 7.6	60.18***	5,908	26
Routine Corrections	4.9	4.0 – 5.9	23.43*	2,944	12
Preselected high risk	11.1	6.8 – 17.6	.00	298	2
Remaining Samples	8.0	6.8 – 9.4	16.90	2,666	12
Between-level Q			**19.85*** **		
Overall	7.1	6.4 – 8.0	86.07***	6,106	27
Routine Corrections	4.9	4.0 – 5.9	23.43*	2,944	12
Preselected high risk	16.6	11.7 – 22.9	6.62	496	3
Remaining Samples	8.0	6.8 – 9.4	16.90	2,666	12
Between-level Q			**39.12*** **		
Routine Corrections					
Overall	7.1	6.4 – 7.8	74.97***	6,106	26
Routine Corrections	4.6	3.7 – 5.6	20.45**	2,502	9
NonRoutine	8.6	7.6 – 9.9	28.86*	3,604	17
Between-level Q			**25.66*** **		

Table continues.

Table 15 continued.

	$B_{0(2)}$ *M*	95% C.I.	*Q*	*n*	*k*
Routine & Treatment Types					
Overall	6.5	5.7 – 7.4	56.72***	5,214	19
Routine Corrections	4.6	3.7 – 5.6	20.45**	2,502	9
Moderate treatment	7.3	6.0 – 8.8	2.21	1,540	5
Mod-High Intensity Rx	9.3	7.2 – 12.0	14.04**	1,172	5
Between-level Q			**20.02*** **		
Preselected					
Overall	7.1	6.4 – 8.0	74.30***	6,016	24
Routine Corrections	4.6	3.7 – 5.6	20.45**	2,502	9
Preselected for treatment	6.9	5.8 – 8.3	4.62	2,011	7
Preselected for risk	11.7	9.1 – 14.9	4.60	1,100	6
Preselected for psychiatric	12.5	8.7 – 17.6	3.08	403	2
Between-level Q			**41.55*** **		
Preselected version 2					
Overall	7.1	6.4 – 8.0	74.30***	6,016	24
Routine Corrections	4.6	3.7 – 5.6	20.45**	2,502	9
Preselected for treatment	6.9	5.8 – 8.3	4.62	2,011	7
Preselected for risk/need	11.9	9.7 – 14.6	7.79	1,503	8
Between-level Q			**41.44*** **		
Preselected version 3					
Overall	7.1	6.4 – 8.0	74.30***	6,016	24
Routine Corrections – Non-US	4.0	3.2 – 5.1	12.33	2,261	7
Routine Corrections – US	8.3	5.1 – 13.4	1.39	241	2
Preselected for treatment	6.9	5.8 – 8.3	4.62	2,011	7
Preselected for risk/need	11.9	9.7 – 14.6	7.79	1,503	8
Between-level Q			**48.17*** **		

Note. Saum was excluded from all analyses.

a. Removed Knight & Thornton (2007) from “two or more.”
b. Removed Bengtson (2008) from Europe
c. Removed Boer (2003) from White.
d. Removed Knight & Thornton (2007) from preselected high risk.

To identify a moderator variable for possible incorporation in the new Static-99 recidivism norms, the variable should ideally demonstrate robust findings (e.g., significant effects in multiple analyses) and an effect that is more than trivial in magnitude (i.e., large enough to be of practical value). The first phase of analyses tested each moderator without controlling for other moderators. Key moderators identified in the first phase were retained for a second stage of analysis that examined their effects after controlling for other key moderators.

Recidivism criteria. In Cox regression, recidivism criteria was significantly related to sexual recidivism (χ^2 change = 4.0, *df* = 1, *p* = .044), with samples using charges (*k* = 11, *n* = 3,6455) showing approximately 1.3 times the recidivism rate of samples using convictions (*k* = 14, *n* = 4,912; Exp(B) = 1.34, 95% C. I. of 1.31 to 1.37). Recidivism criteria also had a significant interaction with Static-99 scores (χ^2 change = 10.5, *df* = 1, *p* = .001).

The meta-analysis included one additional sample (Endrass et al., 2009), although the overall sample size decreased from 8,557 to 6,106. The predicted recidivism rate for a Static-99 score of 2 was 9.3% in samples using charges as their recidivism outcome, and 5.8% in samples using convictions. Variability due to recidivism criteria was significant (between-level *Q* = 16.51, *df* = 1, *p* < .001), and the rate ratio for these findings was 1.6, meaning that the recidivism rate was 1.6 times higher in samples using charges. Although the variability explained by this moderator was significant, there was still significant variability across studies within each level, suggesting that substantial variability is left unaccounted for.

To further understand the interaction between Static-99 and recidivism criteria, the predicted recidivism rates for each Static-99 score were plotted for samples using charges versus samples using convictions (see Figures 11 and 12). Estimates were calculated using logistic regression analyses with Static-99 scores centered on zero. At 5 years, recidivism rates for most Static-99 scores were slightly higher for charges (k = 12, n = 2,336) versus convictions (k = 15, n = 3,949), but the pattern reversed at scores of 9 and higher, with higher recidivism rates for convictions. A similar pattern emerged at 10 years, with a larger difference in recidivism rates between charges (k = 6, n = 783) and convictions (k = 13, n = 1,864), with Static-99 scores of 10 and 11 showing higher recidivism rates for convictions.

The findings from these graphs are uninterpretable because it is not logically possible for charges to yield a *lower* recidivism rate than convictions. This illogical interaction appears to be due to the greater predictive accuracy of Static-99 in samples using convictions as the outcome criteria compared to charges (at 5 years, B_1 = .36 for convictions and .26 for charges; at 10 years, B_1 = .34 for convictions and .23 for charges).

A stronger test of this variable would involve directly comparing charge rates versus conviction rates within the same samples. Five samples provided information on charges and convictions separately, allowing for within-sample comparisons (Bengtson, 2008; Brouillette-Alarie & Proulx, 2008; Cortoni & Nunes, 2007; Epperson, 2003; Johansen, 2007). These samples contained a total of 1,318 offenders, of whom 181 (13.7%) were charged with a new sexual offence, but only 159 (12.1%) were subsequently convicted, meaning that recidivism rates were only 1.1 times

Figure 11. Comparing 5 year sexual recidivism using charges versus convictions

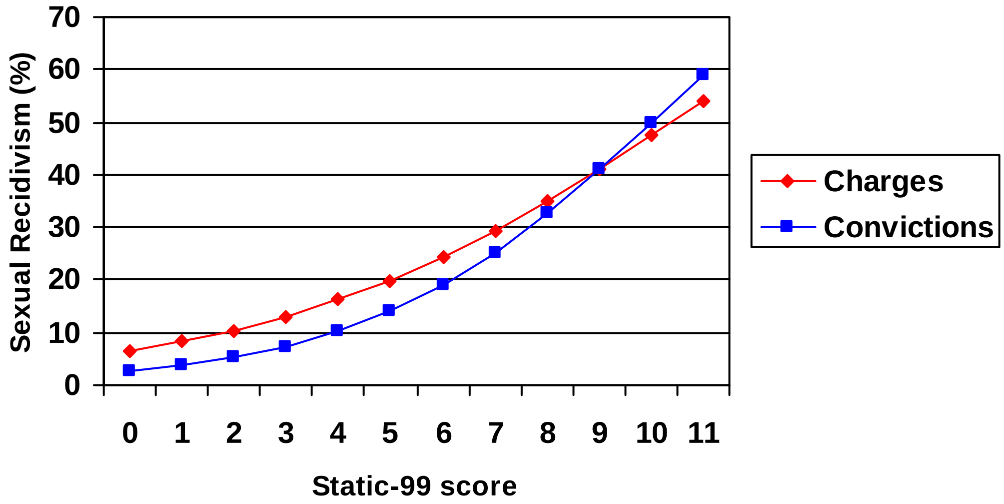
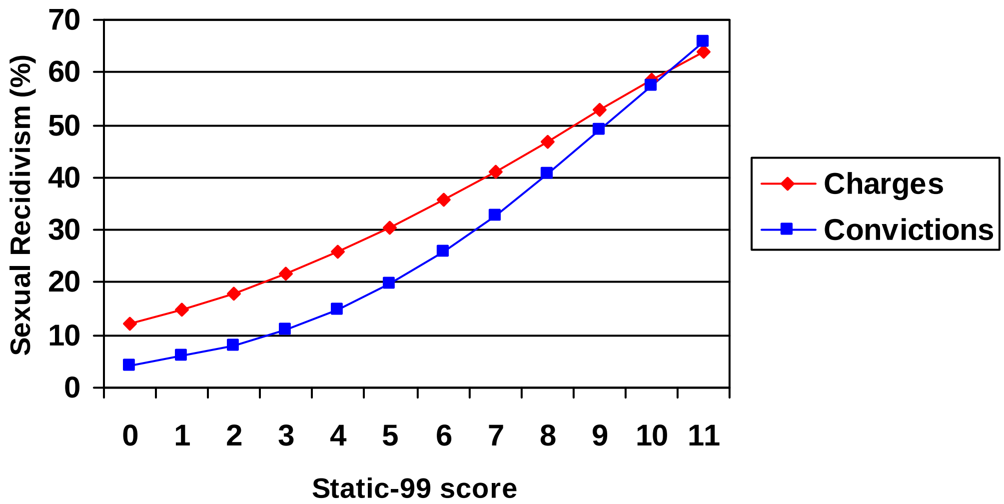


Figure 12. Comparing 10 year sexual recidivism using charges versus convictions



greater for charges versus convictions. The rate ratio (1.1) was notably smaller than that found through Cox regression (1.3) and the logistic regression meta-analysis (1.6).

The Cox regression and meta-analysis findings for this variable were difficult to interpret because they were confounded with other sample-level variables and displayed nonsensical interactions. The within-sample comparisons, although based on smaller sample sizes, was a clearer test. These results indicated small differences, suggesting that the use of charges or convictions as a recidivism definition is unlikely to explain the observed variability in recidivism base rates. Consequently, this variable was not retained for further consideration in the new Static-99 recidivism norms.

Number of recidivism sources used. Most studies used only one source to measure recidivism (in the Cox regression analyses, $k = 17$, $n = 6{,}208$), while 5 studies used two sources ($n = 972$) and 3 studies used four sources ($n = 1{,}377$). Although the number of recidivism sources used was significantly associated with recidivism in both the Cox regression analyses (χ^2 change = 4.7, $df = 1$, $p = .030$) and the meta-analysis (between-level $Q = 10.09$, $df = 2$, $p < .01$), this trend was not obviously linear. From the meta-analysis, samples using two sources had a predicted recidivism rate of 5.6%, which was lower than samples using only one source (6.8%). The significant findings appeared largely due to the unusually high predicted recidivism rate of the three samples using four recidivism sources (predicted recidivism rate of 11.0%). Of these three studies, approximately two thirds of the sample size came from one study (Knight & Thornton, 2007), which was an unusual sample of high-risk sex offenders screened for civil commitment. The other two samples with four recidivism sources (Craissati et al., 2008; Hanson et al., 2007) had base rates no different than studies using fewer recidivism sources.

Due to the small number of studies using more than one recidivism source and the nonlinear findings, this variable was reanalyzed as a dichotomous variable. Both Cox regression and fixed effect meta-analysis found that studies using more than one recidivism source did not have higher base rates than studies using only one source (in Cox regression, χ^2 change = 1.1, $df = 1$, $p = .252$; in the meta-analysis, between-level $Q = 1.84$, $df = 1$, $p > .10$). Interestingly, Knight and Thornton's (2007) sample became a statistical outlier in this analysis, and the remaining seven samples with more than one recidivism source had slightly lower predicted recidivism rates (5.5%) than studies using one recidivism source (6.8%; these rates are from the meta-analysis). This variable was not retained for further consideration in the new Static-99 norms.

Using national criminal records. Using centralized national criminal records to measure recidivism was unrelated to recidivism base rates in both Cox regression (χ^2 change = 0.004, $df = 1$, $p = .951$) and the meta-analysis (between-level $Q = 0.43$, $df = 1$, $p > .75$), although only four studies did not use national records ($n = 1{,}379$ in Cox regression and $n = 1{,}240$ in the meta-analysis). This variable was not retained for further consideration in the new Static-99 norms.

Street time. Street time deducts time spent in prison for non-sexual offences from the follow-up time used for sexual recidivism. One study (Ternowski, 2004) had sufficient information to use street time for the survival analysis variables (used in

Cox regression), but not for the fixed follow-up periods (used in the logistic regression meta-analysis). For Cox regression, four studies used street time (n = 1,501), while 21 studies did not (n = 7,056). Street time was not significantly associated with recidivism (χ^2 change = 1.8, df = 1, p = .179), although the trend was in the expected direction, with recidivism rates 1.1 times higher in studies using street time as opposed to real (i.e., calendar) time.

In the meta-analysis, a smaller sample size using street time was available (k = 3, n = 537) and most of these cases (n = 433) came from Knight and Thornton (2007). Street time was significantly related to recidivism (between-level Q = 14.74, df = 1, p < .001), although the variability across studies in samples that did not use street time was significant (Q = 57.01, df = 22, p < .001), indicating that much variability was still unexplained. Given that 80% of the cases using street time came from an unusually high risk sample, and that these findings conflicted with the Cox regression results (based on a larger sample), this variable was not retained in the new Static-99 recidivism norms.

Cited the coding rules. Citing the Static-99 coding rules (as a proxy measure for assessment quality) was unrelated to recidivism base rates in both Cox regression (χ^2 change = 0.1, df = 1, p = .758) and the meta-analysis (between-level Q = 1.14, df = 1, p > .25). Contrary to the hypothesis, the variability among studies not citing the coding rules was not significant (Q = 2.62, df = 9, p > .90). This variable was not retained for further consideration in the new Static-99 norms.

Provincial versus federal jurisdiction. In Cox regression analyses, offenders in the Canadian provincial correctional systems (k = 4; n = 958) had similar recidivism rates as federal offenders (*k* = 10; *n* = 2,179; χ^2 change = 0.7, *df* = 1, *p* = .414), with slightly lower rates among federal offenders (rate ratio of .903). The meta-analysis results were also non-significant (between-level *Q* = 0.56, *df* = 1, *p* > .25) but in the opposite direction, with federal offenders (*k* = 10, *n* = 1,363) showing higher recidivism rates (predicted recidivism rate of 7.2% compared to 5.9% among provincial offenders). This variable was not retained for further consideration in the new Static-99 norms due to the contradictory and non-significant findings.

Offender type. Cox regression found no significant relationship between offender type and recidivism, both with and without using sample as a strata variable (using sample as strata, χ^2 change = 0.1, *df* = 1, *p* = .800), with rapists (*k* = 12, *n* = 2,182) showing slightly lower sexual recidivism rates than child molesters (*k* = 13, *n* = 3,188). Offender type was significant in the meta-analysis, however (between-level *Q* = 5.05, *df* = 1, *p* < .05), with a predicted recidivism rate of 5.5% among rapists (*k* = 12, *n* = 1,460), compared to 8.0% among the child molesters (*k* = 12, *n* = 2,014). The variability across studies in the recidivism rates of rapists was no more than would be expected by chance (*Q* = 14.81, *df* = 11, *p* > .10), although there was still significant variability across studies for child molesters (*Q* = 30.27, *df* = 11, *p* < .01).

Despite large sample sizes in both analyses, significant results were only found in the meta-analysis. Additionally, the direction of findings is contrary to results found using a similar actuarial measure, Static-2002, where rapists showed higher sexual recidivism rates than child molesters (Hanson et al., in press). This is surprising given that the Static-2002 analyses used five datasets (*n* = 1,860), which were all included in the present analyses. Given that the results were not consistently significant, nor were they consistent with previous findings using similar measures, this variable was not retained for further consideration in the new sexual recidivism norms for Static-99 (although this variable may be re-examined in future analyses of violent recidivism norms).

Country. In Cox regression, country was treated as a categorical variable with the United States used as the reference category. Five samples were from the U.S. (n = 1,784), with eleven samples from Canada (n = 3,252), three from the U.K. (n = 525), five from continental Europe (n = 2,502), and one from New Zealand (n = 493). Country was not significantly related to recidivism (χ^2 change = 6.5, df = 4, p = .164), although all countries generally had lower recidivism rates than the U.S. (except for New Zealand, which was similar to the United States).

A large and significant effect for country was found in the meta-analysis (between-level Q = 27.39, df = 4, p < .001), with the highest predicted recidivism rate in the United States at 8.9% (k = 5, n = 1,516). The one sample from New Zealand showed similar recidivism rates to the United States (8.3%). The predicted recidivism rates were lower in Canada (6.8%; k = 11, n = 1,793) and the United Kingdom (5.4%; k = 3, n = 491), although the confidence intervals from these countries overlapped. With one outlying study removed from continental Europe (Bengtson, 2008), the predicted recidivism rate was 3.8% in Europe, which was significantly lower than both Canada and the United States (non-overlapping confidence intervals). Significant variability across studies was found only in the United States (Q = 14.61, df = 4, p < .01). Although the findings were not consistently significant, the magnitude of the difference between continental Europe and the United States was sufficiently large in the fixed effect meta-analysis to warrant further consideration in the second stage of analyses.

Age at release. Age at release was available for 7,878 cases from 21 samples and was examined using Cox regression with sample entered as strata (analyses are summarized in Table 16). Age at release had a significant negative linear relationship with sexual recidivism (χ^2 change = 28.7, $df = 1$, $p < .001$). The rate ratio was .98 (95% C. I. of .98 to .99), meaning that each one-unit increase in age was associated with 98% of the recidivism rate from the previous (younger) age. In other words, the expected recidivism rate of thirty-two year old offenders is 98% of the recidivism rate of thirty-one year old offenders, which is 98% of the recidivism rate of thirty year old offenders, and so on.

Non-linearity was tested by entering age at release squared in addition to the original age variable (and Static-99 scores). Adding the squared variable was significantly related to recidivism (χ^2 change = 10.7, $df = 1$, $p = .001$), meaning that there is a curvilinear relationship. Adding a cubed age at release variable was non-significant (χ^2 change = 1.5, $df = 1$, $p = .218$), indicating that the line depicting the relationship between age and recidivism does not need two curves. Given the significant findings and the relatively large effect of age at release, it was retained as a moderator variable for the second stage of analyses.

Table 16

Cox Regression Age Analyses

	χ^2 change			Regression coefficient			Rate ratio	
	change	*df*	*p*	*b*	SE	*p*	Exp(B)	95% C.I.
Step 1								
Age	28.67	1	<.001	-.016	.003	<.001	.985	.979 – .990
Step 2								
Age				.040	.018	.024	1.041	1.005 – 1.077
Age^2	10.69	1	.001	-.0007	.0002	.002	.999	.999 – 1.000
Step 3								
Age				.126	.071	.078	1.134	.986 – 1.304
Age^2				-.003	.002	.101	.997	.994 – 1.001
Age^3	1.52	1	.218	-.00002	.00001	.208	1.000	.999990 – 1.000044

Note. The sample size for these analyses was 7,878 from 21 samples, with 957 recidivists. All analyses are also controlling for Static-99 and using original sample as strata.

Year of release. Year of release information was available for 8,556 cases from 25 samples. Initial Cox regression analyses found a strong effect for year of release (χ^2 change = 15.9, *df* = 1, *p* < .001), with each year showing approximately 99% of the recidivism rate from the previous year, with a significant interaction between Static-99 scores and year of release (χ^2 change = 11.9, *df* = 1, *p* = .001). When samples were entered as a strata variable, however, this effect disappeared (χ^2 change = 1.3, *df* = 1, *p* = .247) and switched directions (rate ratio = 1.01), although the interaction with Static-99 scores remained significant (χ^2 change = 11.4, *df* = 1, *p* = .001).

The null findings were surprising given substantial and consistent extant evidence of declining crime rates since the early 1990s. Different distributions of release years among samples, however, suggested that using sample as a strata variable may be too conservative a test (i.e., removing variability due to release year). This was plausible given the dramatic change in results when sample was entered as a strata variable.

The most persuasive test of release year would examine within-sample effects, ideally with large sample sizes having a wide range and even distribution of release years. Of the 25 samples with year of release information, 22 spanned five or more years, and 17 spanned ten or more years. Surprisingly, Cox regression analyses within each of the 25 samples yielded significant results for only 5 samples (and one of them was in the opposite direction; results not reported). These sample-level analyses did not provide persuasive evidence for year of release effects.

Because recidivism rates would likely show complicated (i.e., non-linear) relationships with year of release, the recidivism rates were plotted to allow a visual inspection of the data. Figure 13 depicts the 5 and 10 year recidivism rates for year of release divided into 5-year categories. (The year label on the *x* axis depicts the beginning of the category. For example, 1990 refers to the years 1990-1994). Sample size information for each category is presented in Table 17. Although there is a clear pattern of rising recidivism rates from 1955 to 1974, virtually all of the cases from those years were from one sample (Knight & Thornton, 2007). A drop in recidivism rates is apparent from 1975 to 1989, although extant research suggests the drop should have commenced in the early 1990s, not in the 1970s and 1980s. No clear

Figure 13. Sexual recidivism rates over time (5-year categories)

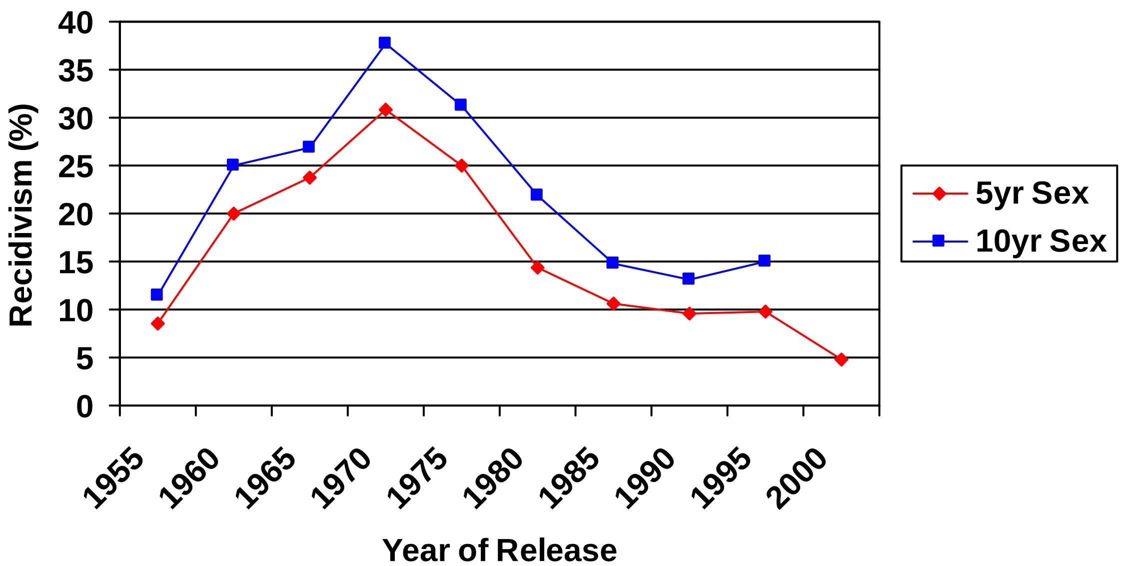


Table 17

Sample Sizes for Year of Release (5-year categories)

	5 year fixed follow-up		10 year fixed follow-up	
Year of Release	Total *n*	Recidivists	Total *n*	Recidivists
1955-1959	35	3	35	4
1960-1964	100	20	96	24
1965-1969	76	18	71	19
1970-1974	133	41	127	48
1975-1979	132	33	77	24
1980-1984	312	45	301	66
1985-1989	664	71	620	92
1990-1994	1,799	172	1,022	134
1995-1999	2,454	238	232	35
2000-2004	310	15		

patterns are discernible after 1990 (the substantial drop in 5 year recidivism rates for offenders released in 2000 or later was based on a small sample size).

Because the largest sample sizes were clustered around the 1990s, 2-year categories were examined to provide a closer look at that time period (see Figure 14 for the plot and Table 18 for the sample sizes). Although random fluctuations are apparent at five years, there seems to be a small trend for lower recidivism rates at 1992 and onwards. At ten years, however, no pattern is readily discernible.

Overall, results for this moderator were not consistently significant. There appeared to be some sort of effect, but the decline in recidivism rates was approximately 15 years earlier than expected (as opposed to peaking in the early 1990s and declining thereafter). Although some evidence was found for this variable, it was insufficient to justify including year of release further as a moderator variable.

Race. Race was examined three different ways, partially due to differences in how studies reported this variable. The three race variables were as follows (sample sizes are for Cox regression): Aboriginal ($k = 7$, $n = 335$) versus non-Aboriginal offenders ($k = 8$, $n = 2,234$), White ($k = 7$, $n = 1,817$) versus all non-White offenders ($k = 7$, $n = 349$), and White ($k = 7$, $n = 1,817$) versus non-White offenders excluding Aboriginals ($k = 5$, $n = 95$). Using sample as a strata variable in Cox regression, the results were not significant for any of the three race variables (for Aboriginal, χ^2 change = 3.1, $df = 1$, $p = .077$; for Non-White, χ^2 change = 3.2, $df = 1$, $p = .073$; for Non-Aboriginal Non-White, χ^2 change = 0.5, $df = 1$, $p = .471$). The meta-analysis results were non-significant for Non-White offenders, both including and excluding

Figure 14. Sexual recidivism rates over time (2-year categories)

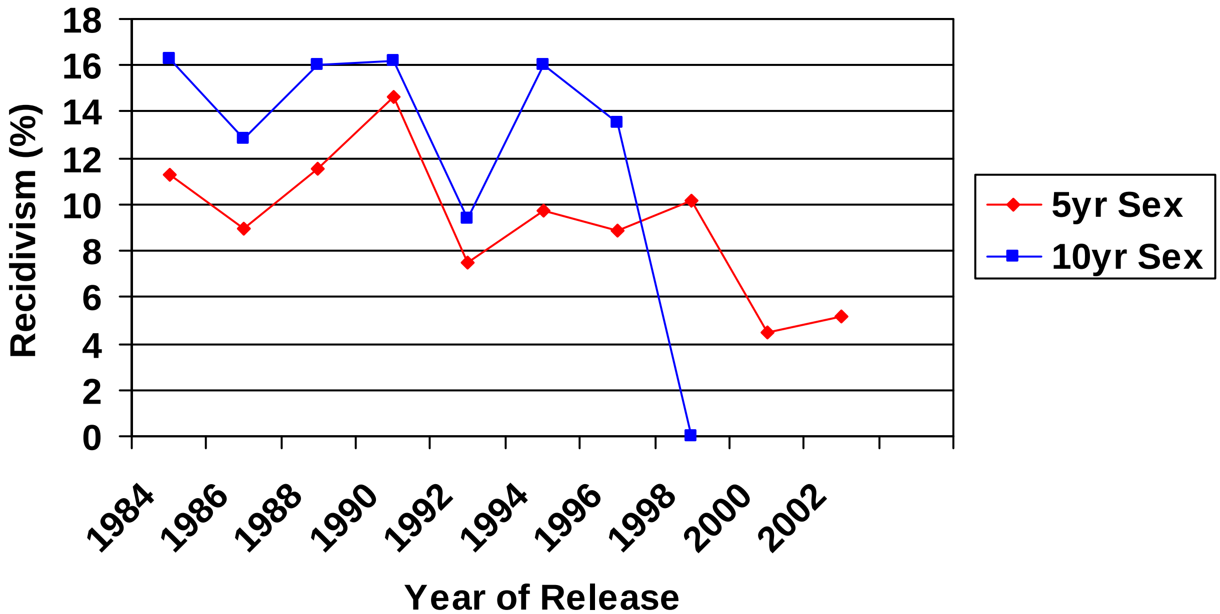


Table 18

Sample Sizes for Year of Release (2-year categories)

	5 year fixed follow-up		10 year fixed follow-up	
Year of Release	Total *n*	Recidivists	Total *n*	Recidivists
1984-1985	203	23	203	33
1986-1987	268	24	266	34
1988-1989	286	33	244	39
1990-1991	446	65	339	55
1992-1993	750	56	458	43
1994-1995	1,450	141	382	61
1996-1997	1,253	112	74	10
1998-1999	354	36	1	0
2000-2001	157	7		
2002-2003	153	8		

Aboriginals (all non-White, between-level $Q = 2.40$, $df = 1$, $p > .10$; for non-White excluding Aboriginals, between-level $Q = 2.16$, $df = 1$, $p > .10$).

The Aboriginal variable, however, was significant (between-level $Q = 4.20$, $df = 1$, $p < .05$), with Aboriginals ($k = 7$, $n = 177$) showing an average predicted recidivism rate of 14.2% for a Static-99 score of 2, compared to 6.8% for non-Aboriginals ($k = 8$, $n = 1,411$). Although this difference was large, the findings were based on a small sample size. Given that the Cox regression analyses (with approximately double the sample size) found no significant results and a much smaller difference (with Aboriginal offenders showing 1.3 times the recidivism rate of non-Aboriginal offenders), this variable was not retained for further consideration in the new sexual recidivism norms for Static-99.

Treatment. Four variables were used to code participation in treatment (the sample sizes presented here are for the Cox regression analyses; see Table 15 for the sample sizes used in the meta-analysis). At the individual level, the variable with the richest detail (but the smallest sample sizes) compared offenders who completed treatment ($k = 5$, $n = 1,065$), dropped out of treatment ($k = 3$, $n = 379$), and those who did not attend treatment ($k = 4$, $n = 537$). Two other dichotomous variables were coded; one compared whether offenders started treatment ($k = 13$, $n = 2,808$) to those who did not ($k = 5$, $n = 436$), and another variable compared offenders who completed treatment ($k = 5$, $n = 1,065$) to offenders who did not complete treatment ($k = 4$, $n = 736$).

One variable was coded at the sample-level, with eight samples considered mostly treated ($n = 2,040$), six samples mixed in their treatment exposure ($n = 1,844$), and only one sample was mostly untreated (n = 1,278). Because only one untreated sample was available, this variable could not be meaningfully examined and was excluded from the moderator analyses.

The three case-level treatment variables were non-significant, both in Cox regression analyses and in the meta-analysis. In Cox regression using sample as strata, for “treatment status,” χ^2 change = 4.2, df = 2, p = .120; for “started treatment,” χ^2 change = 0.2, df = 1, p = .628; for “completed treatment,” χ^2 change = 1.0, df = 1, p = .324; in the meta-analysis, for “treatment status,” between-level Q = 2.96, df = 2, p > .10; for “started treatment,” between-level Q = 0.64, df = 1, p > .25; for “completed treatment,” between-level Q = 2.76, df = 1, p > .05. The treatment status variable and the completed treatment variable were consistently in the expected direction, with the lowest recidivism among treated offenders and the highest among treatment dropouts. The “started treatment” variable was in the expected direction for the meta-analysis, but not in Cox regression, where treatment starters had slightly higher recidivism rates than offenders who did not start treatment (1.07 times higher).

Sample type. This variable was defined in numerous ways. Consistently significant results and the perceived importance of this moderator (evidenced by large effect sizes), as well as previous criticisms of earlier conceptualizations (Abbott, 2009) led to closer examination and several post-hoc refinements to enhance the applied usefulness of this variable. As opposed to providing a step-by-step summary of the evolution of this variable, a brief conceptual overview will be presented instead. For further details, the reader is encouraged to refer to Tables 14 and 15, as well as the supplementary tables in Appendix D.

One initial conceptualization of sample type focused on mental health versus correctional settings (labeled as "setting"). Although samples from mental health settings demonstrated approximately double the recidivism rate of samples managed by correctional systems (in Cox regression, χ^2 change = 60.7, $df = 2$, $p < .001$; in the meta-analysis, between-level $Q = 26.43$, $df = 2$, $p < .001$), further examination of the individual samples, however, revealed that these strong effects were due primarily to two high base rate samples (Bengtson, 2008; Knight & Thornton, 2007). Examining individual studies did not reveal a consistent sample type effect defined by mental health involvement.

The more promising approach focused on the identification of "routine" correctional samples, which was defined as a relatively random (i.e., unselected) sample 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. Simply categorizing samples as routine and non-routine produced a large effect, with routine samples showing a predicted recidivism rate of 4.6%, compared to 8.6% in non-routine samples (between-level $Q = 25.66$, $df = 1$, $p < .001$).

Routine correctional samples thus defined, further efforts to understand the nature of the preselection in the remaining samples yielded two major categories. The first category consisted of offenders preselected on the basis of treatment needs (i.e., through some formal or informal process, offenders were judged as requiring treatment intervention).

The second category consisted of offenders preselected based on a perceived high level of risk and/or need. These samples typically included offenders considered for some rare (i.e., infrequent) measure/intervention/sanction typically reserved for the highest risk cases (e.g., detention until warrant expiry, indefinite sentence). This measure/intervention could involve treatment; samples from high-intensity treatment programs reserved for a small subset of offenders and assigned on the basis of a perceived high level of risk and/or needs were classified in this category (e.g., civil commitment, regional treatment centres in CSC). All other treatment samples were classified in the first category. Offenders preselected for psychiatric systems based on a combination of psychiatric and risk-related reasons were also included in this category.

Two samples were excluded from this categorization. Cortoni and Nunes (2007) consisted of CSC offenders referred for low and moderate intensity treatment, with the low intensity typically reserved for low risk and low need offenders (CSC, 1996). Hill and colleagues (2008) consisted of a representative sample of offenders defined solely by offence severity (sexual homicide perpetrators).

The final sample type variable (labeled "preselected version 2" in the tables) therefore consists of three categories: routine correctional samples, samples preselected on the basis of treatment needs, and samples preselected as high risk/need (including high psychiatric needs). This variable was significant (between-level Q = 41.44, $df = 2$, $p < .001$), with the lowest predicted recidivism rates among routine correctional samples (4.6%; $k = 9$, $n = 2{,}502$). Samples preselected as high risk/need had more than double the recidivism rate of routine samples (11.9%; $k = 8$, n =

1,503), with samples preselected for treatment needs showing intermediate recidivism rates (6.9%; $k = 7$, $n = 2{,}011$). The confidence intervals do not overlap, indicating that each group is significantly different from each other. Additionally, the variability across samples among the preselected for treatment need samples and the preselected as high risk/need samples was no more than would be expected by chance.

There was, however, significant variability among the routine correctional samples ($Q = 20.45$, $df = 8$, $p < .01$). One routine sample (Bartosh et al., 2003) had a noticeably higher base rate (11.2%) than all other samples, whose base rates were no higher than 6.4%. This difference may have been due to jurisdiction, as Bartosh and colleagues (2003) was an American sample and previous analyses suggested that base rates may be higher in the U.S. After removing Bartosh et al. (2003) and the only other routine American sample (Epperson, 2003), there was no longer significant variability across samples among the routine category ($Q = 12.33$, $df = 6$, $p > .05$). When the routine American samples were removed, the predicted recidivism rate for routine correctional samples changed from 4.6% to 4.0% (the predicted recidivism among the two routine American samples was 8.3%, which interestingly was higher than the predicted recidivism rate of samples preselected on the basis of treatment needs).

These analyses may suggest that routine correctional samples show consistently low recidivism base rates only outside the U.S. This should be interpreted with caution, however, given that there were only two routine American samples and although one displayed an unusually high base rate compared to the other routine correctional samples, the other American sample (Epperson, 2003) appeared roughly similar to the non-American routine samples. The sample type variable (keeping the American samples in the routine category) was retained for further analyses in the new recidivism norms for Static-99.

The Final Moderator Variables for the New Static-99 Norms

Of all the moderator variables examined, the only three that warranted further consideration were country, age at release, and sample type. These variables were entered together in a logistic regression analysis at 5 and 10 years, after controlling for Static-99. At 5 years (see Table 19), all three moderators contributed significantly to the prediction of recidivism. For sample type (Wald statistic = 57.71, $df = 2$, $p < .001$), both the routine correctional samples and the samples preselected based on treatment needs had significantly lower recidivism rates than the samples preselected based on high risk/need (odds ratios of .41 and .54, respectively). For country (Wald statistic = 12.68, $df = 4$, $p = .013$), only Canada had significantly lower recidivism rates than the United States (in Canada, the odds of recidivism was .75 of the odds of recidivism in the U.S.). For age, a significant negative curvilinear relationship was found (for age, Wald statistic = 3.92, $df = 1$, $p = .048$; for age squared, Wald statistic = 7.74, $df = 1$, $p = .005$).

At 10 years (see Table 20), only age at release and sample type remained significant predictors of recidivism. For sample type (Wald statistic = 37.86, $df = 2$, $p < .001$), both the routine correctional samples and the samples preselected based on treatment needs had significantly lower recidivism rates than the samples preselected based on high risk/need (odds ratios of .35 and .64, respectively). For age, a

Table 19

Five Year Logistic Regression Results Using Three Main Moderator Variables

	b	SE	Wald	*df*	*p*	Exp(B)	95% C.I.
Static-99	.285	.021	179.63	1	<.001	1.33	1.28 – 1.39
Sample Type (risk)			57.71	2	<.001		
routine	-.890	.125	50.98	1	<.001	.41	.32 – .52
treatment	-.614	.129	22.57	1	<.001	.54	.42 – .70
Country (U.S.)			12.68	4	.013		
Canada	-.287	.118	5.90	1	.015	.75	.60 – .95
UK	.069	.207	0.11	1	.739	1.07	.71 – 1.61
Europe	-.244	.141	2.97	1	.085	.78	.59 – 1.03
New Zealand	.398	.214	3.46	1	.063	1.49	.98 – 2.27
Age	.051	.026	3.92	1	.048	1.05	1.00 – 1.11
Age Squared	-.001	<.001	7.74	1	.005	.999	.998 – 1.00
Constant	-3.076	.510	36.35	1	<.001	.05	

Note. For categorical variables, the level in parentheses was used as the reference category.

Table 20

Ten Year Logistic Regression Results Using Three Main Moderator Variables

	b	SE	Wald	*df*	*p*	Exp(B)	95% C.I.
Static-99	.299	.028	116.00	1	<.001	1.35	1.28 – 1.42
Sample Type (risk)			37.86	2	<.001		
routine	-1.060	.175	36.70	1	<.001	.35	.25 – .49
treatment	-.445	.159	7.84	1	.005	.64	.47 – .88
Country (US)			6.24	4	.182		
Canada	-.288	.183	2.49	1	.115	.75	.52 – 1.07
UK	.209	.244	0.74	1	.391	1.23	.76 – 1.99
Europe	.056	.162	0.12	1	.728	1.06	.77 – 1.45
New Zealand	.629	.526	1.43	1	.231	1.88	.67 – 5.26
Age	.071	.032	5.01	1	.025	1.07	1.01 – 1.14
Age Squared	-.001	<.001	6.45	1	.011	.999	.998 – 1.00
Constant	-3.378	.636	28.17	1	<.001	.034	

Note. For categorical variables, the level in parentheses was used as the reference category significant negative curvilinear relationship was found (for age, Wald statistic = 5.01, *df* = 1, *p* = .025; for age squared, Wald statistic = 6.45, *df* = 1, *p* = .011). Country was not a significant predictor of 10 year recidivism rates after controlling for age and sample type (Wald statistic = 6.24, *df* = 4, *p* = .182).

Meta-regression was used as a more powerful test of the moderators. In meta-regression, linear regression is used to analyze effect sizes across samples (Hedges, 1994). Age could not be incorporated in this analyses given that it was a case-level variable. Effect sizes were predicted using each study's weight (the inverse of the variance) as the WLS weight variable in SPSS. Dummy-coded variables for sample type and country were used as independent variables, with sample type entered first. The statistic to test the contribution of country was calculated by hand using Formula 27 of Hedges (1994). The change in *Q* from adding country was not significant (*Q* = 6.98, *df* = 3, *p* > .05), meaning that once sample type was controlled for, country did not add significantly to the prediction of sexual recidivism.

Given that the meta-regression and the 10 year logistic regression results both found that differences across countries were no longer significant after controlling for age and sample type, country was dropped from the analyses. The new Static-99 recidivism norms should therefore incorporate only age and sample type.

Age at Release: Developing the Static-99R

The age at release analyses reported earlier found a curvilinear effect of age. Further analyses were therefore conducted to examine the shape of the distribution. Figure 15 presents the contribution of age and age squared to the prediction of sexual recidivism, estimated from Cox regression. The hazard ratio in the *y* axis is centered on a Static-99 score of 0. Figure 16 presents the same analyses using logistic regression at 5 years, with Static-99 scores centered on 2 and the *y* axis referring to recidivism rates as a percentage. Both figures demonstrate the same pattern, with a slight increase in risk for offenders in their twenties, followed by a fairly linear decrease in risk up until their eighties.

The relationship between age and recidivism, as depicted in Figures 15 and 16, may be slightly misleading, however, given that offenders less than 25 years old are already given a risk point on Static-99. To better understand the relationship between age and recidivism, Static-99 was re-computed without the age item. For offenders with Static-99 total scores but no individual item scores, the age-free Static-99 could be computed if age information was available elsewhere in the dataset; if not, it was coded as missing. Figure 17 presents the same analysis as Figure 16, but with the age-free Static-99 score. When the age item is excluded from Static-99, the relationship between age and recidivism is clearly linear. Similar results were found in Cox regression, where age added significantly to the prediction of sexual recidivism after controlling for the age-free Static-99 score (χ^2 change = 48.8, $df = 1$, $p < .001$), but the squared age value did not (χ^2 change = 2.9, $df = 1$, $p = .087$).

Rather than fitting a complicated curvilinear age adjustment to the original Static-99 scale, a simpler option would be to develop a new age item. For this task, the overall sample was divided into two subsamples: one for the development of new Static-99 age weights, and another sample for validation. Of offenders with age-free Static-99 scores and information on age at release ($k = 23$, $n = 8{,}128$), offenders with

Figure 15. Cox regression hazard ratios as a function of age at release

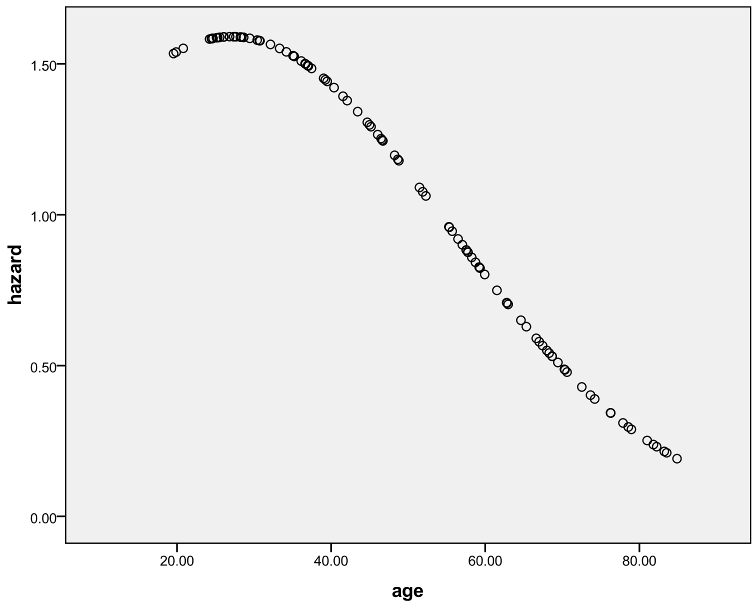


Figure 16. Recidivism as a function of age at release (using logistic regression)

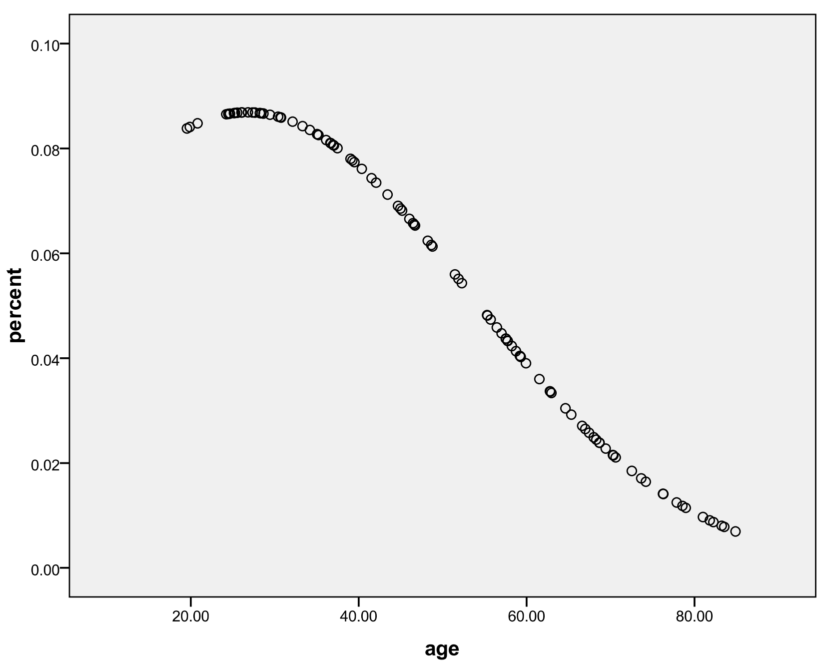
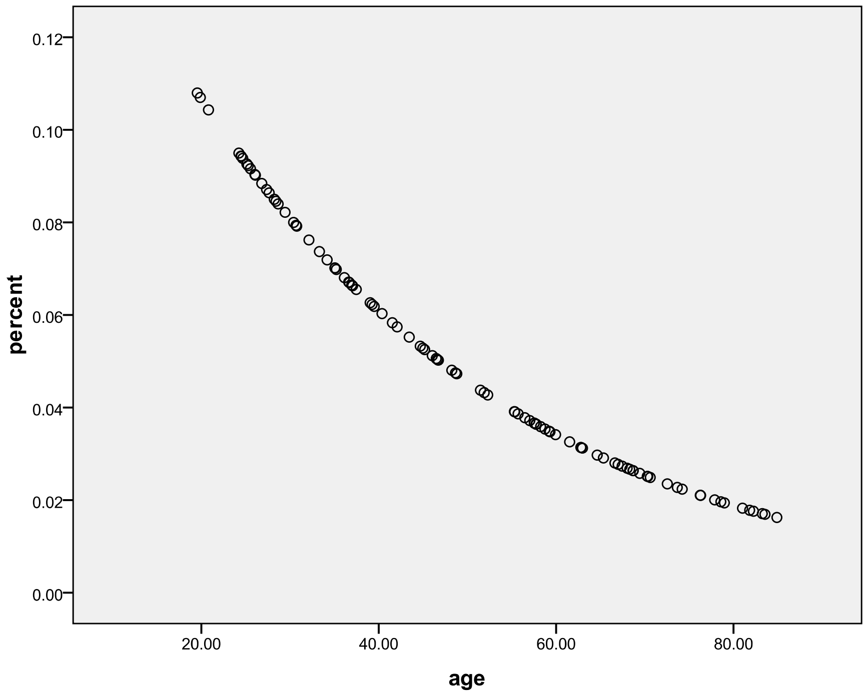


Figure 17. Recidivism as a function of age at release (using logistic regression and removing the age item from Static-99)



follow-up periods of less than 10 years (*k* = 23, *n* = 5,736) were used as the development sample, while offenders with follow-up periods greater than 10 years were retained for validation (*k* = 15, *n* = 2,392). This method allowed the new Static-99 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.

The principles guiding the selection of the new age weights were as follows: a) each unit should approximate the Static-99 units found in other analyses (rate ratio and odds ratio of approximately 1.35); b) offenders with the median age (39 years old) should receive a score of 0 for this item; c) the revised measures should have higher overall predictive accuracy than the original; and d) age should no longer contribute significantly once the Static-99 scale with the new age item is included.

The relationship between age and recidivism in the construction sample (*n* = 5,736) was explored by Karl Hanson and David Thornton (the developers of Static-99) and the new age weight they proposed was as follows: offenders less than 35 would receive 1 point on Static-99, offenders age 35 to 39.999 would receive 0 points, offenders age 40 to 59.999 would have 1 point subtracted from Static-99, and offenders age 60 and older would have 3 points subtracted from Static-99. The revised Static-99 scale with these age weights will be referred to as Static-99R.

Comparing Static-99R to Static-99 in the validation sample, there was a slight increase in relative predictive accuracy for Static-99R, as measured by the Area Under the Receiver Operating Characteristic Curve (using fixed follow-up periods: at 5 years, ROC for Static-99R was .720, compared to .713 for Static-99; at 10 years,

Table 21

Comparing Static-99R to Static-99

	Static-99R					Static-99				
	χ^2 change	*df*	*p*	Exp(B)	95% C.I.	χ^2 change	*df*	*p*	Exp(B)	95% C.I.
Logistic Regression 5 years										
Static scale	135.82	1	<.001	1.34	1.27 – 1.41	128.92	1	<.001	1.37	1.29 – 1.44
Age	1.27	1	.260			4.14	1	.042		
Logistic Regression 10 years										
Static scale	164.45	1	<.001	1.32	1.27 – 1.39	157.91	1	<.001	1.36	1.29 – 1.42
Age	1.46	1	.227			5.54	1	.019		
Cox Regression										
Static scale	180.14	1	<.001	1.28	1.23 – 1.32	165.92	1	<.001	1.29	1.24 – 1.34
Age	0.66	1	.418			7.87	1	.005		

ROC for Static-99R was .710, compared to .706 for Static-99). Table 21 displays the results of logistic regression analyses (at 5 and 10 years) and Cox regression analyses, all controlling for routine versus non-routine samples. In all analyses, age at release did not add significant predictive accuracy after controlling for Static-99R, whereas it did add incremental predictive accuracy to the original Static-99. These results indicate that the original Static-99 did not sufficiently account for age at release, whereas the revised scale does.

No statistical shrinkage (whereby effect sizes tend to decrease in independent replications) was observed with Static-99R; in fact, in two of three analyses, the effect of Static-99R increased in the validation sample (5 year fixed ROC = .709 in the development sample compared to .720 in the validation sample; in the 5-year logistic regression analyses, the odds ratio was 1.32 in the development sample versus 1.34 in the validation sample; in the Cox regression analyses, the rate ratio was 1.36 in the development sample versus 1.28 in the validation sample).

Static-99R Recidivism Rates

The meta-analytic findings from the Static-99R logistic regression analyses at 5 and 10 years are presented in Table 22, with B_0 centered on a Static-99R score of 2. For the 5 and 10 year findings from the individual samples, see Appendixes E and F, respectively. Note that from the comparable Static-99 tables presented earlier (Tables 11 & 12), five additional samples were excluded due to insufficient age information to code Static-99R (Craig et al., 2006; de Vogel et al., 2004; Endrass et al., 2009; Langton, 2003; Milton, 2003). As in the moderator analyses, Saum (2007) was excluded as an outlier. The findings were aggregated using both random effect and

Table 22

Meta-Analysis of Static-99R Logistic Regression Coefficients with B_0 Centered on 2

	Fixed Effect			Random Effect			
	M	95 % CI	*Q*	*M*	95 % CI	*k*	*n*
Five Years							
B_1	.290	.251 – .329	18.98	.290	.251 – .329	20	5,543
$B_{0(2)}$	7.4%	6.6 – 8.3	64.27***	6.9	5.4 – 8.8	22	5,591
Ten Years							
B_1	.262	.211 – .312	14.64	.270	.210 –.330	13	2,373
$B_{0(2)}$	12.3%	10.7 – 14.1	47.31***	10.8	7.8 – 14.9	13	2,373

Note. Total sample size and studies differ between B_0 and B_1 because B_1 cannot be computed without recidivists.

fixed effect meta-analysis. If *Q* was less than degrees of freedom, the random effect analyses were the same as the fixed effect findings. When they differed, the random effect analyses were relied on because it is more conservative, it explicitly acknowledges variability across samples in the estimates, and the intent of these analyses was to generalize to a population (Hedges & Vevea, 1998; Schmidt, Oh, & Hayes, 2009).

The predictive accuracy (B_1) of Static-99R was similar to Static-99 (at 5 years, B_1 = .290 using both random and fixed effects; at 10 years, B_1 = .262 using fixed effect, and B_1 = .270 using random effect). The 5 year predicted recidivism rate for a Static-99R score of 2 was 7.4% using fixed effect analyses and 6.9% using random effect analyses. At 10 years, the predicted recidivism rate for a score of 2 was 12.3%

using fixed effect analyses and 10.8% using random effect analyses. Similar to Static-99, the variability in predictive accuracy (B_1) across samples was no more than would be expected by chance, but there was significant variability across samples in the recidivism base rates ($B_{0(2)}$).

With Static-99R, sample type showed a strong effect in explaining base rate variability, similar to Static-99 (meta-analytic results are presented in Table 23). At 5 years, the predicted recidivism rate for a Static-99R score of 2 was 5.0% for routine correctional samples, 7.2% for samples preselected based on treatment need, and 12.2% for samples preselected as high risk/need (between-level Q = 34.78, df = 2, p < .001). The confidence intervals for the group preselected as high risk/need did not overlap with the confidence intervals for the other two groups. Similar to Static-99, there was significant variability across studies in the routine correctional samples (Q = 19.57, df = 7, p < .01), but not in the samples preselected based on treatment need or as high risk/need. The routine correctional sample was the only group where the Q value was greater than the degrees of freedom. Random effect meta-analysis produced the same predicted recidivism rate (5.0%), although the confidence intervals were wider.

At 10 years, the predicted recidivism rate for a Static-99R score of 2 was 5.1% for routine correctional samples, 11.5% for samples preselected based on treatment need, and 19.7% for samples preselected as high risk/need (between-level Q = 42.57, df = 2, p < .001). At 10 years, the variability across studies was no more than would be expected by chance in all three groups.

Table 23

Static-99R Logistic Regression Meta-Analysis Across Sample Type, With $B_{0(2)}$ Converted to a Percentage

	$B_{0(2)}$ *M*	95% C.I.	*Q*	*n*	*k*
5 years					
Overall	7.4	6.6 – 8.3	62.83***	5,501	20
Routine Corrections	5.0	4.1 – 6.2	19.57**	2,406	8
Routine – Random Effect	5.0	3.2 – 7.8			
Preselected for treatment	7.2	6.0 – 8.8	4.47	1,782	6
Preselected high risk/need	12.2	9.9 – 15.0	4.01	1,313	6
Between-level Q			**34.78*** **		
10 years					
Overall	12.3	10.7 – 14.1	47.20***	2,319	12
Routine Corrections	5.1	3.5 – 7.3	3.37	750	4
Preselected for treatment	11.5	9.4 – 14.0	0.59	866	5
Preselected high risk/need	19.7	15.9 – 24.1	0.67	703	3
Between-level Q			**42.57*** **		

One option for incorporating base rate variability when reporting Static-99R scores is to report the estimates separately for the three sample types. As an example of what such recidivism norms would look like, estimated Static-99R recidivism rates were calculated using the logistic regression coefficients from the random effect meta-analysis. Figures 18 and 19 present the 5 and 10 year estimated recidivism rates for each of the three sample types (a table with the recidivism estimates is included in Appendix G). Appendix H presents the equations used in these calculations (in the form of SPSS syntax). Given that the predictive accuracy (B_1) was stable (the variability across samples was non-significant), the combined B_1 from all three

Figure 18. Static-99R 5 year sexual recidivism rates

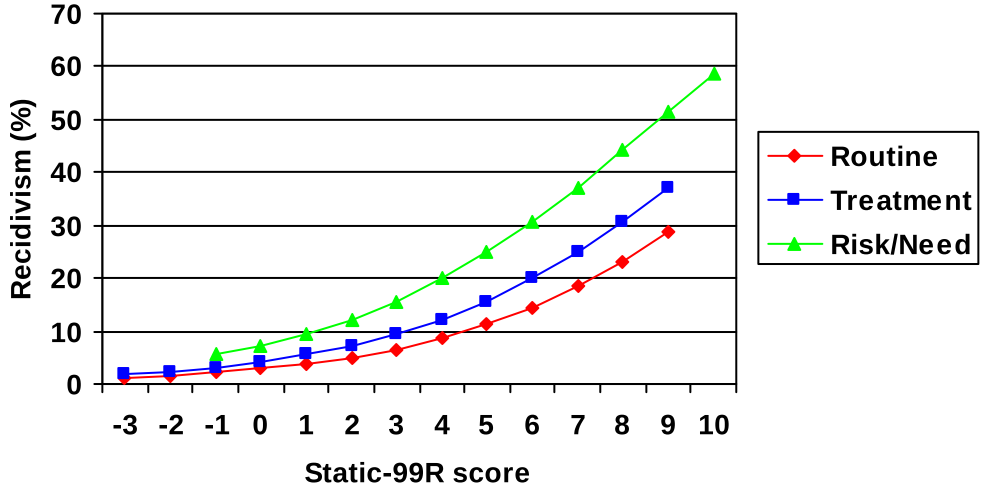
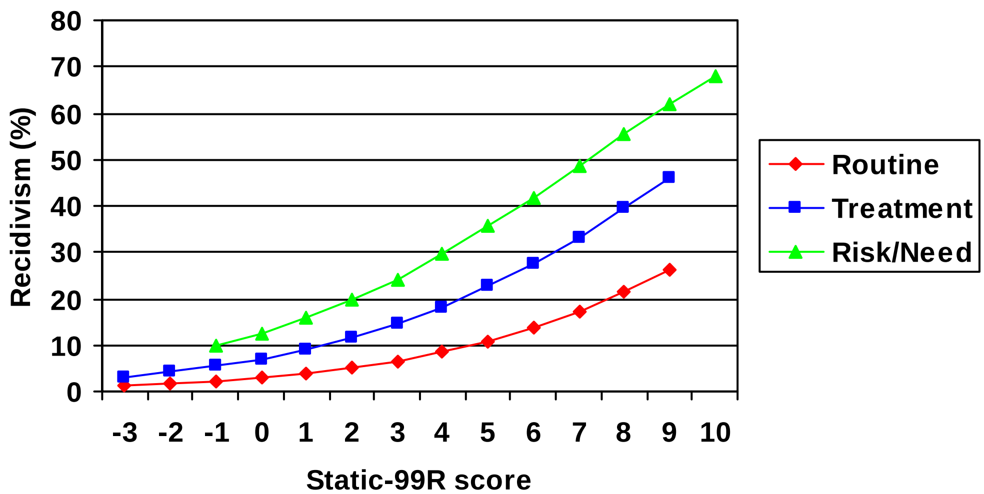


Figure 19. Static-99R 10 year sexual recidivism rates



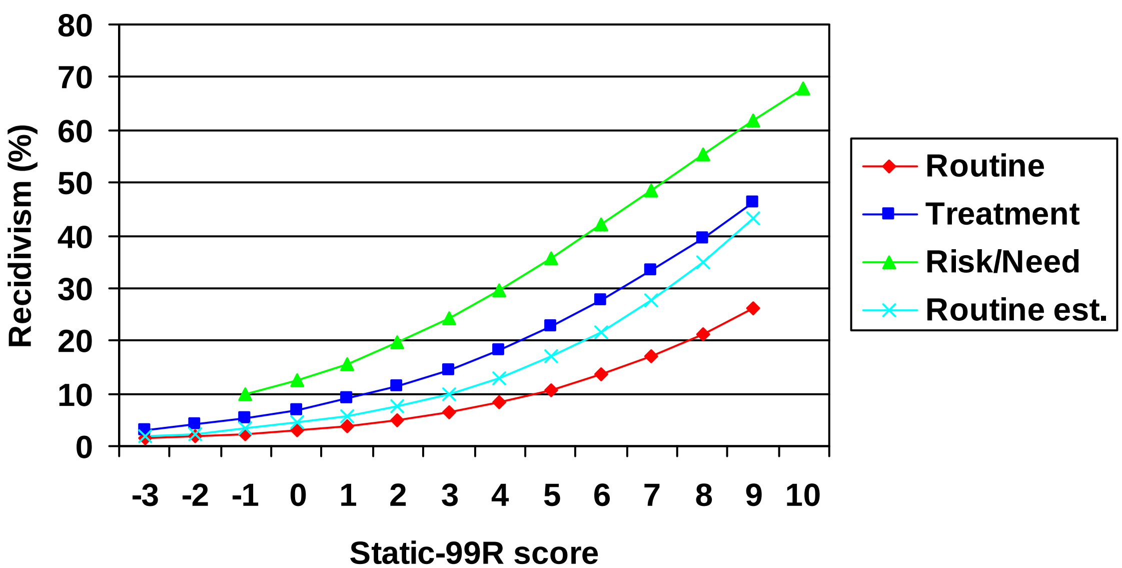
sample types was used. To avoid extrapolating to Static-99R scores not represented in the data, the estimates were truncated to include only scores with a minimum of 8 cases available at 5 years.

For Static-99R scores of 4 through 9 in the routine correctional samples, the recidivism estimates are slightly lower at 10 years than at 5 years. Although the difference is small (the largest difference is for scores of 9, with 28.7% at 5 years and 26.3% at 10 years), this finding is logically impossible given that recidivism is cumulative. The anomalous findings occurred because the sample size used at 10 years (k = 4, n = 750) is approximately one third the sample size used at 5 years (k = 8, n = 2,406), and over 85% of the 10-year routine cases were from two samples (Boer, 2003; Langström, 2004) with somewhat lower base rates relative to the other routine samples.

Although the finding of lower recidivism rates at 10 years compared to 5 years among routine samples is explainable, the anomaly nonetheless creates a problem for applied risk evaluators. One possible solution is to abstain from reporting 10-year recidivism estimates for routine correctional samples on the grounds that insufficient samples were available to provide an estimate that was reliable and plausible (i.e., higher than the 5 year estimate).

Another solution is to adjust the 10-year estimates to what would be expected if more cases had sufficient follow-up information (this is the basic premise of survival analysis and Cox regression). One way to do this is to multiply the 5-year estimates by a rate ratio. Survival analysis of all routine correctional cases (n = 4,041) generated a rate ratio of 1.51, meaning that the 10 year rate was 1.51 times higher than the 5 year rate (the estimated recidivism rate was 6.7% at 5 years and 10.2% at 10 years). Figure 20 presents the 10 year recidivism estimates for all three sample types (the same as Figure 19), but adds another line, which consists of the 5-year estimates for routine samples multiplied by 1.51 (the values from the graph are also presented in Appendix G).

Figure 20. Static-99R 10 year sexual recidivism rates with adjustment to routine samples



Discussion

Although many analyses were conducted, three major findings emerged from the current study. The first is that Static-99 provides consistent estimates of relative risk. The second is that recidivism base rates per Static-99 score are significantly and meaningfully lower in more recent samples compared to the original Static-99 recidivism norms. The third major finding is that base rates per Static-99 score are not consistent across samples. Further examination of this variability identified two important moderator variables: age at release and sample type.

In all meta-analyses (examining both Static-99 and Static-99R at 5 and 10 years), the variability across studies in relative predictive accuracy was no more than would be expected by chance. Each one-unit increase on Static-99 or Static-99R was associated with an increase of approximately 1.35 in the odds of recidivism.

Previous meta-analyses consistently found significant variability in the predictive accuracy of Static-99 (Hanson & Morton-Bourgon, 2004, 2007, 2009), but these analyses used Cohen's *d* as the measure of effect size. Cohen's *d* is based on the same statistical model as ROC curves (Swets, 1986) and is therefore subject to the same limitations, namely that the effect size can be strongly influenced by variability in the distribution of predictor values (Static-99 scores; Hanson, 2008). Using logistic regression, which is a more stable measure of relative predictive accuracy (Hanson, 2008), no significant variability was found.

The stability in relative predictive accuracy is impressive given the large number of studies and the diversity of samples included (from child molesters in Allan et al., 2007, to sexual homicide perpetrators in Hill et al., 2008). Evaluators can therefore use Static-99 for relative risk decisions (e.g., resource allocation) with a reasonable level of confidence.

The second major finding from the current study was that recidivism base rates per Static-99 score are meaningfully lower in current samples. In other words, the original Static-99 recidivism norms are over-estimating recidivism. It is unclear whether recidivism rates have decreased since the offenders from the original samples were released or whether the original samples simply had abnormally high recidivism rates. A combination of both options is also possible. Given substantial extant evidence of declining crime rates since the early 1990s (Bureau of Justice Statistics, 2006; Federal Bureau of Investigation, 2007; Finkelhor & Jones, 2006; Jones & Finkelhor, 2006; Minnesota Department of Corrections, 2007; Mishra & Lalumière, 2009a; Mishra & Lalumière, 2009b; Public Safety Canada, 2008), a reduction in the recidivism rates likely plays some part in the current findings. It is worth noting that although recidivism rates (repeat offending) are not the same as crime rates (all offending, including repeat and first-time), they are similar phenomena and given that crime rates include recidivistic behaviour, trends in recidivism rates should roughly mirror trends in crime rates.

The question that naturally follows from these findings is why the recidivism rates would be lower in current samples. It is unlikely that a single factor is responsible for these trends. Possible explanations that have been proposed include demographic factors (e.g., aging population, increased obesity, reliance on medications such as Prozac or other serotonin-affecting agents), cultural factors (e.g., changing mores regarding sexuality, increased awareness about sexual assault leading to greater vigilance and supervision of children), and criminal justice system factors (e.g., offender treatment, increased supervision, deterrent/incapacitation effects of longer sentences—for a summary, see Finkelhor & Jones, 2006).

Although understanding the causes of the lower recidivism rates is both interesting and important, it is not necessary for examining how these changes should affect best practices in offender assessment, management, and supervision. In other words, even without understanding the reasons for the change, the evidence of change forces evaluators to adjust their practice. Documenting a decline in recidivism rates is therefore a distinct task from explaining that decline, and further research in both areas is needed.

The last major finding of the current study is that although the relative predictive accuracy of Static-99 was consistent across studies, the predicted recidivism rates for each Static-99 score showed significant variability. Of 13 moderator variables examined, only 2 showed consistently significant results in accounting for the observed base rate variability: age at release and sample type.

The original, dichotomous Static-99 age item (was the offender less than 25 years old) was insufficient in accounting for the negative linear relationship between age at release and sexual recidivism. The age item was modified and given new weights, and the resulting scale was called Static-99R. The new age weighting in Static-99R is consistent with empirical evidence that current actuarial scales do not sufficiently incorporate age at release (Barbaree, Langton, & Blanchard, 2007; Hanson, 2006). Findings that age information did not significantly add to the prediction of sexual recidivism after controlling for Static-99R suggests that evaluators using this revised scale do not need to consider age at release further in their assessment.

The other notable moderator was sample type, which examined the extent to which some samples were preselected and how this affected recidivism rates, after controlling for Static-99 or Static-99R. The lowest base rates were found for routine, unselected correctional samples, with the highest base rates among samples preselected as high risk/need, and intermediate base rates among samples selected as having a need for treatment. These differences were non-trivial in magnitude and should therefore have some impact on applied risk evaluations. Although sample type was a systems-level factor, it appeared to identify offenders who differed on individual characteristics that might affect how they are processed through the correctional system (such as perceived risk and treatment need).

The findings for age and sample type suggest that base rate differences are related to differences in offender characteristics more than differences in methodological factors or other sample-level factors, such as country or year of release. The lack of significant findings for methodological and systems variables is a promising indicator of the generalizability of Static-99. The strong effect for individual factors, however, highlights that Static-99 does not consider all relevant risk factors, and some of the unmeasured risk factors are clearly important.

The large differences across sample types and their potential importance for applied decision-making necessitate further efforts to understand the nature of the preselection that occurred in the samples used for the current study. The very fact that preselected samples showed higher recidivism rates after controlling for Static-99 (or

Static-99R) means that offenders must have been preselected on the basis of risk-relevant factors external to Static-99. Given the consistency with which the diverse selection procedures applied in the different studies increased the base rate, it is relatively easy to preselect higher base rate samples. The two different levels of preselection (based on treatment need and high risk/need) indicate that the more preselection on risk-relevant criteria that occurs, the higher the recidivism base rate.

The variability in recidivism rates among the different levels of preselection reflects the minimum range of effect that factors external to Static-99 (but internal to the individual) can have. The observed range does not define bounds on the full effect of factors external to Static-99 because the exact nature of the preselection process would have differed across samples, with some preselection processes more strongly related to risk than others. It may therefore be possible to preselect a group with even higher base rates than the estimates from the samples identified as high risk/need.

Given that detailed information on what factors were considered when offenders were placed in various settings is not available in the majority of current studies, any conceptual understanding of the preselection processes is therefore based on logical inferences and not empirical data. Such conceptual inferences about the preselection may, nonetheless, be useful in understanding how relatively informal selection procedures (e.g., referring offenders to certain agencies or treatment programs) could produce such large selection effects.

Conceptually, routine samples of sex offenders contain the full distribution of unspecified risk relevant characteristics that exists in the population from which the sample is drawn. Offenders with low levels of risk and need (on factors external to

Static-99) are less likely to be referred to treatment programs or for exceptional measures (e.g., civil commitment, detention until warrant expiry, specialized psychiatric assessments). Although this would particularly be the case in jurisdictions where access to treatment and other resources is scarce, virtually all settings have some prioritization of resources, with some offenders not targeted. Samples with moderate levels of preselection (e.g., referred for treatment) may therefore represent a truncated range of the "routine" sample, with many of the low risk/need offenders screened out.

Whereas the first level of preselection screens out the lowest risk offenders, the higher level of preselection (identifying high risk/need offenders) focuses on the highest risk offenders. In these samples, the low risk/need offenders would almost be completely excluded and offenders with moderate risk/need would be somewhat underrepresented. The two preselected groups identified (the treatment need group and the high risk/need group) may represent different types of filtering, with the first group merely excluding the lowest risk offenders, and the second group targeting the highest risk offenders.

Even assuming that the correctional systems are not particularly accurate in these preselection processes (e.g., a sizable minority of low risk offenders are still referred for treatment), relatively small selection effects that disproportionately screen out low risk cases (or screen in high risk cases) can produce the observed variation in base rates, provided that at least some risk relevant characteristics external to Static-99 were considered in the selection process. Many of the decision-making processes in the criminal justice system would plausibly consider (implicitly or explicitly) some individual risk factors such as sexual deviance, procriminal attitudes, continued antisocial behaviour in prison, and this would produce notable selection effects. This conceptualization therefore implies that by referring offenders for certain settings/resources/measures, correctional systems can easily produce samples with higher base rates than routine correctional samples.

Strengths and Limitations of the Current Study

The main strengths of this study were the large sample sizes and the collection of diverse studies. The large sample sizes permitted numerous high-power analyses, and the use of multiple samples provided an opportunity for analyses generally not possible in a single study (e.g., comparing different types of samples, examining age with large groups of elderly offenders). Other strengths were the strict inclusion criteria for samples and cases, and the extensive data cleaning, which helped restrict the analyses to samples with complete and appropriate data.

The primary limitations of the current study relate to the restrictions inherent in using data from prior studies designed for different purposes than the current investigation. Some variables may be coded inconsistently across studies (e.g., offender type), and the quality of data coding is often difficult to determine based on the data provided (although the data were cleaned, only some types of errors could be detected without access to the original files).

Although the large sample sizes used in the current study would increase the possibility of a Type I error, the use of multiple analyses (e.g., Cox regression, logistic regression meta-analysis) and conservative decision criteria (e.g., requiring large effects) limited the potential impact of such errors. The number of post-hoc analyses, however, should invoke some caution in interpreting the findings, although once again, the conservative inclusion of moderator variables somewhat mitigates this concern.

Implications for Applied Assessment

The findings from the current study indicate that Static-99 (and Static-99R) is a robust measure of relative risk, but using any absolute recidivism estimate requires justification. The three sets of norms developed in the current study, though demonstrating large effects, were developed post-hoc. These norms are more plausible than the original estimates simply because they are based on complete data from samples that are larger, more representative, and more recent. Although more plausible, these norms should not be adopted in all contexts without caution. Given that base rates showed such large variability across samples, evaluators cannot take for granted that any set of norms is going to apply to the context in which they are assessing risk.

The variability across samples indicates that evaluators concerned about absolute recidivism rates cannot easily make reliable judgements about absolute risk based solely on Static-99 or Static-99R. General base rates and risk factors external to the measures should be incorporated to maximize the prediction of absolute recidivism rates. One option is for correctional systems to develop their own sets of norms. Developing customized estimates based on recent, local samples may give the evaluator more confidence that the norms are applicable to their sample. Appendix H contains the equations needed for any jurisdiction to input their own base rate (in the form of a logistic regression coefficient) and calculate Static-99R recidivism estimates. For jurisdictions or contexts unable to generate their own Static-99R recidivism estimates, the best solution may be to report the norms for all three sample types and note that they reflect the possible effects of variability due to unmeasured risk factors. Another alternative could be to use the routine, unselected samples because they are the most representative of the population of all adjudicated sex offenders.

Implications for Risk Assessment

These findings indicate that actuarial prediction (using Static-99 or Static-99R) does not include all risk-relevant factors, and the unmeasured risk factors produce meaningful variation in base rates. Dawes and colleagues (1989) have previously noted that actuarial scales should be continuously re-evaluated and revised to reflect increases in knowledge. The current findings should therefore be used to improve actuarial measures. For example, in the development of the original Static-99, there were an insufficient number of offenders over the age of 60 to evaluate their recidivism risk relative to younger offenders (personal communication with Karl Hanson, August 25, 2009). A larger sample size of offenders over 60 in the current study (n = 600) allowed for an item revision that sufficiently incorporates age at release (resulting in the Static-99R). The current study is therefore part of a research agenda that can contribute to the continuous growth of actuarial risk assessments by incorporating new risk factors.

Although this study extends previous research on factors incrementally predictive to static actuarial scores (Barbaree, Langton, & Blanchard, 2007; Beech et al., 2002; Dempster & Hart, 2002; Hanson, 2006; Hanson et al., 2007; Marques, Wiederanders, Day, Nelson, & van Ommeren, 2005; Olver et al., 2007; Thornton, 2002), further research is still needed. Although this study evaluated 13 moderator variables, there were some potential moderators for which insufficient information was available to test (e.g., community supervision, dynamic risk factors, detection rates) or for which more detailed analyses may yield stronger results (e.g., treatment or year of release). Future studies can extend this research by replicating the current findings and prospectively testing more moderators. Further research is also needed to provide violent and general recidivism estimates for Static-99R (with a re-examination of offender type as a moderator variable) and to assess whether the cut-off scores for the relative risk categories for Static-99 (i.e., low, moderate-low, moderate-high, and high) should be revised for Static-99R.

References

Note: References marked with an asterisk were included in the analyses. Wilson et al. (2007a & b) both have an asterisk, although they were combined into one sample for the current study.

Abbott, B. R. (2009). Applicability of the new Static-99 experience tables in sexually violent predator risk assessments. *Sexual Offender Treatment, 4*(1), 1-24. Retrieved June 26, 2009 from http://www.sexual-offender-treatment.org/1-2009_03.html

*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.

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

Andrews, D. A., & Bonta, J. (2006). *The psychology of criminal conduct* (4th ed.). Newark, NJ: LexisNexus/Anderson.

Andrews, D. A., Bonta, J. L., & Wormith, J. S. (2004). *Level of Service / Case Management Inventory: An offender assessment system*. Toronto, ON: Multi-Health Systems.

Andrews, D. A., Bonta, J., & Wormith, J. S. (2006). The recent past and near future of risk and/or need assessment. *Crime & Delinquency, 52*(1), 7-27.

Andrews, D. A., Zinger, I., Hoge, R. D., Bonta, J., Gendreau, P., & Cullen, F. (1990). Does correctional treatment work? A clinically relevant and psychologically informed meta-analysis. *Criminology, 28*(3), 369-404.

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.

Babchishin, K., & Hanson, R. K. (2009). Improving our talk: Moving beyond the “Low,” “Moderate,” and “High” typology of risk communication. *Crime Scene, 16*(1), 11-14.

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.

*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.

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.

Beggs, S. M., & Grace, R. C. (2005, November). *Improving sex offender risk assessments: The validity of incorporating measures of treatment outcome.*

Paper presented at the annual Research and Treatment Conference of the Association for the Treatment of Sexual Abusers, Salt Lake City, UT.

*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.

Besserer, S., & Trainor, C. (2000). *Criminal victimization in Canada, 1999 (Juristat. Catalogue no. 85-002-XIE).* Ottawa: Canadian Center for Justice Statistics.

*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).

*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.

Boer, D. P., Wilson, R. J., Gauthier, C. M., & Hart, S. D. (1997). Assessing risk of sexual violence: guidelines for clinical practice. In C. D. Webster & M. A. Jackson (Eds.), *Impulsivity: Theory, assessment, and treatment* (pp. 326-342). New York: Guilford Press.

Bonta, J. (1996). Risk-needs assessment and treatment. In A. T. Harland (Ed.), *Choosing correctional options that work: Defining the demand and evaluating the supply* (pp. 18-32). Thousand Oaks, California: Sage Publications.

Bonta, J. (2002). Offender risk assessment: Guidelines for selection and use. *Criminal Justice & Behavior, 29*(4), 355-379.

Bonta, J., Bogue, B., Crowley, M., & Motiuk, L. (2001). Implementing offender classification systems: Lessons learned. In G. A. Bernfeld, D. P. Farrington, & A. W. Leschied (Eds.), *Offender rehabilitation in practice: Implementing and evaluating effective programs* (pp. 227-246). Chichester, England: John Wiley & Sons.

Bonta, J., LaPrairie, C., & Wallace-Capretta, S. (1997). Risk prediction and re-offending: Aboriginal and non-aboriginal offenders. *Canadian Journal of Criminology, 37* (2), 127-144.

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.

Bonta, J., Rugge, T., Scott, T.-L., Bourgon, G., & Yessine, A. K. (2008). Exploring the black box of community supervision. *Journal of Offender Rehabilitation, 47*, 248-270.

*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.

Brennan, S., & Taylor-Butts, A. (2008). *Sexual assault in Canada: 2004 and 2007.* Ottawa, ON: Statistics Canada, Canadian Centre for Justice Statistics.

Broadhurst, R. G., & Loh, N. S. (1995). Rearrest probabilities for the 1984-1993 apprehended Western Australia population: A survival analysis. *Journal of Quantitative Criminology, 11*, 289-313.

*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.

Brown, J. A. (2003). A comparison of actuarial methods of predicting sexual dangerousness. *Dissertations Abstracts International, 55* (02), 1068B. (UMI No. 3120894).

Bureau of Justice Statistics (2006). *Criminal victimization, 2006*. Retrieved October 9, 2008, from http://www.ojp.usdoj.gov/bjs/abstract/cv06.htm

Cohen, J. (1988). *Statistical power analysis for the behavioural sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Collaborative Outcome Data Committee. (2007a). *Sexual offender treatment outcome research: CODC Guidelines for evaluation Part 1: Introduction and overview* (Corrections Research User Report No. 2007-02). Ottawa, Ontario: Public Safety Canada.

Collaborative Outcome Data Committee. (2007b). *The Collaborative Outcome Data Committee's guidelines for the evaluation of sexual offender treatment outcome research Part 2: CODC guidelines* (Corrections Research User Report No. 2007-03). Ottawa, Ontario: Public Safety Canada.

Correctional Service of Canada. (1996). Standards and guidelines for the provision of services to sex offenders. Retrieved July 30, 2009, from http://198.103.98.138/text/pblct/so/standards/stande-eng.shtml

*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., Beech, A. R., & Browne, K. D. (2006). Evaluating the predictive accuracy of sex offender risk assessment measures on UK samples: A cross-validation of the Risk Matrix 2000 scales. *Sexual Offender Treatment, 1*(1). Retrieved January 23, 2007, from http://www.sexual-offender-treatment.org/19.98.html

*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.

Craissati, J., Webb, L., & Keen, S. (2005). *Personality disordered sex offenders.* Unpublished report. London, UK: Bracton Centre (Oxleas NHS Trust), London Probation Area and Home Office.

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

*de Vogel, V., de Ruiter, C., van Beek, D., & Mead, G. (2004). Predictive validity of the SVR-20 and Static-99 in a Dutch sample of treated sex offenders. *Law and Human Behavior, 28*, 235-251.

Dempster, R. J., & Hart, S. D. (2002). The relative utility of fixed and variable risk factors in discriminating sexual recidivists and nonrecidivists. *Sexual Abuse: A Journal of Research and Treatment, 14*, 121-138.

Doren, D. M. (2002). *Evaluating sex offenders: A manual for civil commitments and beyond*. Thousand Oaks, CA: Sage.

Doren, D. M. (2004). Stability of the interpretative risk percentages for the RRASOR and Static-99. *Sexual Abuse: A Journal of Research and Treatment, 16,* 25-36.

Ducro, C., & Pham, T. (2006). Evaluation of the SORAG and the Static-99 on Belgian sex offenders committed to a forensic facility. *Sexual Abuse: A Journal of Research and Treatment, 18*, 15-26.

Eisenhart, C. (1947). Inverse sine transformation of proportions. In C. Eisenhart, M. W. Hastay, & W. A. Wallis, (Eds.), *Selected techniques of statistical analysis for scientific and industrial research and production and management engineering* (pp. 395-416). New York: McGraw-Hill.

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

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.

Ellis, L., & Peterson, J. (1996). Crime and religion: An international comparison among thirteen industrial nations. *Personality and Individual Differences, 20*, 761-768.

Elmore, J. G., & Gigerenzer, G. (2005). Benign breast disease: The risk of communicating risk. *The New England Journal of Medicine, 353*, 297-299.

*Endrass, J., Urbaniok, F., Held, L., Vetter, S., & Rossegger, A. (2009). The accuracy of the Static-99 in predicting recidivism in Switzerland. *International Journal of Offender Therapy and Comparative Criminology, 53, 482-490.*

*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.

Evans, J. St. B. T., Handley, S. J., Perham, N., Over, D. E., & Thompson, V. A. (2000). Frequency versus probability formats in statistical word problems. *Cognition, 77*, 197-213.

Federal Bureau of Investigation (2007). *Crime in the United States, 2007*. Retrieved October 9, 2008, from http://www.fbi.gov/ucr/cius2007/index.html

Finkelhor, D., & Jones, L. (2006). Why have child maltreatment and child victimization declined? *Journal of Social Issues, 62,* 685-716.

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

Flores, A. W., Lowenkamp, C. T., Holsinger, A. M., & Latessa, E. J. (2006). Predicting outcome with the Level of Service Inventory-Revised: The importance of implementation integrity. *Journal of Criminal Justice, 34*, 523-529.

Friendship, C., Mann, R. E., & Beech, A. R. (2003). Evaluation of a national prison-based treatment program for sexual offenders in England and Wales. *Journal of Interpersonal Violence, 18*, 744-759.

Friendship, C., Thornton, D., Erikson, M., & Beech, A. (2001). Reconviction: A critique and comparison of two main data sources in England and Wales. *Legal and Criminological Psychology, 6*, 121-129.

Furby, L., Weinrott, M. R., Blackshaw, L. (1989). Sex offender recidivism: A review. *Psychological Bulletin, 105*, 3-30.

Gallagher, C. A., Wilson, D. B., Hirschfield, P., Coggeshall, M. B., & MacKenzie, D. L. (1999). A quantitative review of the effects of sex offender treatment on sexual reoffending. *Corrections Management Quarterly, 3*, 19-29.

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*(1), 19-30.

*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.

Hall, G. C. N. (1995). Sexual offender recidivism revisited: A meta-analysis of recent treatment studies. *Journal of Consulting and Clinical Psychology, 63*, 802-809.

Hanson, R. K. (1997). *The development of a brief actuarial risk scale for sexual offense recidivism* (User Report 1997-04). Ottawa, ON: Department of the Solicitor General of Canada.

Hanson, R. K. (2005). Twenty years of progress in violence risk assessment. *Journal of Interpersonal Violence, 20*(2), 212-217.

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

Hanson, R. K. (2008). What statistics should we use to report predictive accuracy. *Crime Scene, 15*(1), 15-17.

Hanson, R. K., Bourgon, G., Helmus, L., & Hodgson, S. (2009). The principles of effective correctional treatment also apply to sexual offenders: A meta-analysis. *Criminal Justice & Behavior, 36*, 865-891.

Hanson, R. K., & Bussière, M. T. (1998). Predicting relapse: A meta-analysis of sexual offender recidivism studies. *Journal of Consulting and Clinical Psychology, 66*(2), 348-362.

Hanson, R. K., Gordon, A., Harris, A. J. R., Marques, J. K., Murphy, W., Quinsey, V. L., & Seto, M. C. (2002). First report of the Collaborative Outcome Data Project on the effectiveness of psychological treatment of sex offenders. *Sexual Abuse: A Journal of Research and Treatment, 14*, 169-194.

Hanson, R. K., & Harris, A. J. R. (2000). Where should we intervene? Dynamic predictors of sexual offense recidivism. *Criminal Justice and Behavior, 27*(1), 6-35.

*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, ON, Canada: Public Safety Canada.

Hanson, R. K., Helmus, L., & Thornton, D. (in press). Predicting recidivism among sexual offenders: A multi-site study of Static-2002. *Law and Human Behavior.*

Hanson, R. K., Lloyd, C. D., Helmus, L., & Thornton, D. (2008). *Using multiple samples to estimate percentile ranks and relative risk ratios for actuarial risk tools: A Canadian example using Static-99 and Static-2002.* Unpublished manuscript.

Hanson, R. K., & Morton-Bourgon, K. E. (2004). *Predictors of sexual recidivism: An updated meta-analysis* (Corrections Research User Report No. 2004-02). Ottawa, ON: Public Safety and Emergency Preparedness Canada.

Hanson, R. K., & Morton-Bourgon, K. E. (2005). The characteristics of persistent sexual offenders: A meta-analysis of recidivism studies. *Journal of Consulting and Clinical Psychology, 73*, 1154-1163.

Hanson, R. K., & Morton-Bourgon, K. E. (2007). *The accuracy of recidivism risk assessments for sexual offenders: A meta-analysis* (Corrections Research User Report No. 2007-01). Ottawa, ON: Public Safety and Emergency Preparedness Canada.

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.

Hanson, R. K., & Nicholaichuk, T. (2000). A cautionary note regarding Nicholaichuk et al. (2000). *Sexual Abuse: A Journal of Research and Treatment, 12*, 289-293.

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

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

Harris, A. J. R., & Hanson, R. K. (2004). *Sex offender recidivism: A simple question* (User Report 2004-03). Ottawa, ON: Public Safety and Emergency Preparedness Canada.

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

Harris, G. T., Rice, M. E., & Hilton, N. Z. (2009, June). *The actuarial assessment and communication of domestic violence risk*. Paper presented at the Annual Convention of the Canadian Psychological Association, Montreal, QC, Canada.

Harris, G. T., Rice, M. E., & Quinsey, V. L. (1998). Appraisal and management of risk in sexual aggression: Implications for criminal justice policy. *Psychology, Public Policy, and Law, 4,* 73-115.

*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.

Hedges, L. V. (1994). Fixed effects models. In H. Cooper & L. V. Hedges, (Eds.), *The handbook of research synthesis* (pp. 285-299). New York: Russell Sage Foundation.

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis.* New York: Academic Press.

Hedges, L. V., & Vevea, J. L. (1998). Fixed- and random-effects models in meta-analysis. *Psychological Methods, 3*, 486-504.

Helmus, L. (2007). *A multi-site comparison of the validity and utility of the Static-99 and Static-2002 for risk assessment with sexual offenders.* Unpublished Honour's thesis, Carleton University, Ottawa, ON, Canada.

Helmus, L. M. D., & Hanson, R. K. (2007). Predictive validity of the Static-99 and Static-2002 for sex offenders on community supervision. *Sexual Offender Treatment, 2*(2), 1-14. Retrieved June 1, 2008 from http://www.sexual-offender-treatment.org/60.0.html

Helmus, L., Hanson, R. K., & Morton-Bourgon, K. E. (in press). International comparisons of the validity of actuarial risk tools for sexual offenders, with a focus on Static-99. In D. P. Boer, L. A. Craig, R. Eher, M. H. Miner, & F. Pfafflin (Eds.), *International perspectives on the assessment and treatment of sexual offenders: Theory, practice, and research*. Wiley.

Helmus, L., Hanson, R. K., & Thornton, D. (2009). Reporting Static-99 in light of new research on recidivism norms. *The Forum, 21*(1), 38-45.

*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.

Hilton, N. Z., Carter, A., Harris, G. T., Sharpe, A. J. B. (2008). Does using nonumerical terms to describe risk aid violence risk communication? *Journal of Interpersonal Violence, 23*, 171-188.

Hood, R., Shute, S., Feilzer, M., & Wilcox, A. (2002). Sex offenders emerging from long-term imprisonment: A study of their long-term reconviction rates and of parole board members' judgments of their risk. *British Journal of Criminology, 42*, 371-394.

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

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.

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

Jones, L. M., & Finkelhor, D. (2003). Putting together evidence on declining trends in sexual abuse: A complex puzzle. *Child Abuse & Neglect, 27*, 133-135.

Karelitz, T. M., & Budescu, D. V. (2004) .You say "probable" and I say "likely": Improving interpersonal communication with verbal probability phrases. *Journal of Experimental Psychology, 10*, 25-41.

Kenworthy, T., Adams, C. E., Brooks-Gordon, B., & Fenton, M. (2004). *Psychological interventions for those who have sexually offended or are at risk of offending* (CD004858). Cochrane Database of Systematic Reviews, Issue 3. Chichester, UK: John Wiley & Sons.

Kilpatrick, D. G., Edwards, C. N., & Seymour, A. E. (1992). *Rape in America: A report to the nation.* Arlington, VA: National Crime Victims Center.

*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.

Krohn, M. D. (2001). A Durkheimian analysis of international crime rates. *Social Forces, 57*, 654-670.

Kutchinsky, B. (1991). *Pornography* and rape: Theory and practice? Evidence from crime data in four countries where pornography is easily available. *International Journal of Law and Psychiatry, 14*, 47-64.

Langan, P. A., Schmitt, E. L., & Durose, M. R. (2003). *Recidivism of sex offenders released from prison in 1994.* Washington, DC: U.S. Department of Justice, Bureau of Justice Statistics

*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.

*Langton, C. M. (2003). Contrasting approaches to risk assessment with adult male sexual offenders: An evaluation of recidivism prediction schemes and the utility of supplementary clinical information for enhancing predictive accuracy. *Dissertations Abstracts International, 64* (04), 1907B. (UMI No. NQ78052).

Laudon, K. C. (1986). Data quality and due process in large interorganizational record systems. *Communications of the ACM, 29*(1), 4-11.

Lösel, F., & Schmucker, M. (2005). The effectiveness of treatment for sexual offenders: A comprehensive meta-analysis. *Journal of Experimental Criminology, 1,* 117-146.

Maffei, S., & Merzagora Betsos, I. (2007). Crime and criminal policy in Italy: Tradition and modernity in a troubled country. *European Journal of Criminology, 4*, 461-482.

Marenin, O. (1997). Victimization surveys and the accuracy and reliability of official crime data in developing countries. *Journal of Criminal Justice, 25*, 463-475.

Marques, J. K., Wiederanders, M., Day, D. M., Nelson, C., & van Ommeren, A. (2005). Effects of a relapse prevention program on sexual recidivism: Final results from California's sex offender treatment and evaluation project (SOTEP). *Sexual Abuse: Journal of Research and Treatment, 17*, 79-107.

Mauer, M. (2006). *Race to incarcerate* (Rev. Ed.). New York: New Press.

Meehl, P. E. (1954). *Clinical versus statistical prediction: A theoretical analysis and a review of the evidence.* Minneapolis: University of Minnesota Press.

McGrath, R. J., Cumming, G. F., & Burchard, B. L. (2003). *Current practices and trends in sexual abuser management: The Safer Society 2002 nationwide survey*. Brandon, VT: Safer Society Press.

McGrath, R. J., Hoke, S. E., Livingston, J. A., & Cumming, G. F. (2001, November). *The Vermont Assessment of Sex-Offender Risk (VASOR): An initial reliability and validity study*. Presentation at the annual Research and Treatment Conference of the Association for the Treatment of Sexual Abusers, San Antonio, TX.

*Milton, J. (2003). *Care pathways and outcome of mentally disordered sex offenders referred to high secure psychiatric hospitals*. Unpublished M.D. thesis, University of Nottingham, Nottingham, United Kingdom.

Minnesota Department of Corrections (2007). *Sex offender recidivism in Minnesota: April 2007*. St. Paul, MN: Author. Retrieved January 9, 2009 from http://www.doc.state.mn.us/publications/documents/04-07SexOffenderReport-Recidivism.pdf

Mishra, S., & Lalumière, M. (2009a). The big drop in sex crimes. *The Forum, 21*(2), 41-48.

Mishra, S., & Lalumière, M. (2009b). Is the crime drop of the 1990s in Canada and the USA associated with a general decline in risky and health-related behavior? *Social Science & Medicine, 68*, 39-48.

Monahan, J., & Silver, E. (2003). Judicial decision thresholds for violence risk management. *International Journal for Forensic Mental Health, 2*, 1-6.

Montana, S., & Thompson, G. (2005, November). *Risk assessment for priest sex offenders: Developing a risk assessment tool to measure relapse risk*. Paper presented at the annual Research and Treatment Conference of the Association for the Treatment of Sexual Abusers, Salt Lake City, UT.

Moore, R. A., Derry, S., McQuay, H. J., & Paling, J. (2008). What do we know about communicating risk? A brief review and suggestion for contextualising serious, but rare, risk, and the example of cox-2 selective and non-selective NSAIDS. *Arthritis Research and Therapy, 10*.

Mossman, D. (1994). Assessing predictions of violence: Being accurate about accuracy. *Journal of Consulting and Clinical Psychology, 62*(4), 783-792.

*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.

Nunes, K. L., Wexler, A., Firestone, P., & Bradford, J. M. (2003, October). *Incarceration and recidivism in a sample of sexual offenders*. Paper presented at the annual Research and Treatment Conference of the Association for the Treatment of Sexual Abusers, St. Louis, MO.

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.

Paolucci, E. O., Genuis, M. L., & Violato, C. (2001). A meta-analysis of the published
research on the effects of child sexual abuse. *The Journal of Psychology, 135*, 17-36.

Public Safety Canada (2008). *Corrections and conditional release statistical overview: Annual report 2008*. Ottawa, ON: Author. Retrieved June 22, 2009 from http://www.publicsafety.gc.ca/res/cor/rep/2008-04-ccrso-eng.aspx

Quinsey, V. L., Harris, G. T., Rice, M. E., & Cormier, C. A. (2006). *Violent offenders: Appraising and managing risk* (2nd Ed.). Washington, DC: American Psychological Association.

Raj, A., Silverman, J. G., Wingood, G. M., & Diclemente, R. J. (1999). Prevalence and correlates of relationship abuse among a community-based sample of low-income African American women. *Violence Against Women, 5*, 272-291.

Resick, P. A. (1993). The psychological impact of rape. *Journal of Interpersonal Violence, 8*, 223-255.

Rice, M. E., & Harris, G. T. (2003). The size and sign of treatment effects in sex offender therapy. *Annals of the New York Academy of Sciences, 989*, 428-440.

Rice, M. E., & Harris, G. T. (2005). Comparing effect sizes in follow-up studies: ROC
area, Cohen's *d*, and *r*. *Law and Human Behavior, 29*(5), 615-620.

Rushton, J. P. (1995). Race and crime: International data for 1989-1990. *Psychological Reports, 76*, 307-312.

*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. *Dissertations Abstracts International, 68* (03), B. (UMI No. 3255539).

Schmidt, F. L., Oh, I.-S., & Hayes, T. L. (2009). Fixed- versus random-effects models in meta-analysis: Model properties and an empirical comparison of differences in results. *British Journal of Mathematical and Statistical Psychology, 62*, 97-128.

Schmucker, M., & Lösel, F. (2008). Does sexual offender treatment work? A systematic review of outcome evaluations. *Psicotherma, 20,* 10-19.

Simpson, K. (2009, June). *The effects of study quality on meta-analytic findings: An exploration of the effectiveness of community supervision*. Presented at the Annual Convention of the Canadian Psychological Association, Montreal, QC, Canada.

Smith, P., Goggin, C., & Gendreau, P. (2002). *The effects of prison sentences and intermediate sanctions on recidivism: General effects and individual differences.* (User Report No. 2002-01). Ottawa, ON: Solicitor General Canada.

Stalans, L. J., Seng, M., Yarnold, P., Lavery, T., & Swartz, J. (2001). Process and initial impact evaluation of the Cook County adult probation department's sex offender program: Final summary report for the period of June, 1997 to June,

2000. Chicago, IL: Loyola University. Retrieved June 23, 2009 from http://www.icjia.state.il.us/public/pdf/researchreports/An%20Implementation_Project%20in%20Cook%20County.pdf

Swets, J. A. (1986). Indices of discrimination and diagnostic accuracy: Their ROCs and implied models. *Psychological Bulletin, 99*, 100-117.

*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).

Terry, K. J. (2008). Stained glass: The nature and scope of child sexual abuse in the Catholic church. *Criminal Justice and Behavior, 35,* 549-569.

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

Tough, S. E. (2001). Validation of two standardized risk assessments (RRASOR, 1997; Static-99, 1999) on a sample of adult males who are developmentally disabled with significant cognitive deficits. *Masters Abstracts International, 39* (06), 1626B. (UMI No. MQ58817).

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*, 1124-1131.

United Nations. (2007*). Responses by indicator to: Questionnaire for the ninth United Nations survey of crime trends and operations of criminal justice systems, covering the period 2003-2004*. United Nations, Office of Drugs and Crime. Retrieved March 29, 2009 from http://www.unodc.org/documents/data-and-analysis/CTS9_by_%20indicator.pdf

Webster, C. D., Douglas, K. S., Eaves, D. & Hart, S. D. (1997). Assessing risk of violence to others. In C. D. Webster & M. A. Jackson (Eds.), *Impulsivity: Theory, assessment, and treatment* (pp. 251-277). New York: Guilford Press.

*Wilson, R. J., Cortoni, F., & Vermani, M. (2007a). *Circles of support and accountability: A national replication of outcome findings* (Report No. R-185). Ottawa, ON: 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. *The Howard Journal, 46*, 327-337.

Appendixes

Appendix A

STATIC-99 Coding Form

Question Number	Risk Factor	Codes	Score
1	Young (S9909)	Aged 25 or older Aged 18 – 24.99	0 1
2	Ever Lived With (S9910)	Ever lived with lover for at least two years? Yes No	 0 1
3	Index non-sexual violence - Any Convictions (S9904)	No Yes	0 1
4	Prior non-sexual violence - Any Convictions (S9905)	No Yes	0 1
5	Prior Sex Offences (S9901)	Charges / Convictions None / None 1-2 / 1 3-5 / 2-3 6+ / 4+	 0 1 2 3
6	Prior sentencing dates (excluding index) (S9902)	3 or less 4 or more	0 1
7	Any convictions for non-contact sex offences (S9903)	No Yes	0 1
8	Any Unrelated Victims (S9906)	No Yes	0 1
9	Any Stranger Victims (S9907)	No Yes	0 1
10	Any Male Victims (S9908)	No Yes	0 1
	Total Score	**Add up scores from individual risk factors**	

TRANSLATING STATIC 99 SCORES INTO RISK CATEGORIES

Score	Label for Risk Category
0,1	Low
2,3	Moderate-Low
4,5	Moderate-High
6 plus	High

Appendix B

Static-99 logistic regression analyses at 5 years

	Static-99 *M* (*SD*)		*N*	*N* recid (%)	B_1	B_1SE	$B_{0(0)}$	$B_{0(0)}SE$	$B_{0(2)}$	$B_{0(2)}SE$	$B_{0(5)}$	$B_{0(5)}SE$
Allan et al. (2007)	2.29	1.98	299	35 (11.7)	0.411	0.086	-3.219	0.357	-2.397	0.229	-1.164	0.230
Bartosh et al. (2007)	3.23	2.16	90	12 (13.3)	0.138	0.135	-2.352	0.589	-2.076	0.388	-1.660	0.357
Bengtson (2008)	3.75	2.09	310	61 (19.7)	0.210	0.070	-2.249	0.333	-1.830	0.216	-1.201	0.154
Bigras (2007)	2.54	1.91	207	19 (9.2)	0.362	0.117	-3.412	0.493	-2.688	0.314	-1.603	0.295
Boer (2003)	3.26	2.29	299	11 (3.7)	0.515	0.146	-5.603	0.892	-4.572	0.627	-3.027	0.335
Bonta & Yessine (2005)	5.21	1.98	101	19 (18.8)	0.265	0.139	-2.924	0.849	-2.394	0.590	-1.600	0.283
Brouillette-Alarie & Proulx (2008)	3.79	2.21	199	29 (14.6)	0.360	0.095	-3.349	0.513	-2.629	0.348	-1.549	0.211
Cortoni & Nunes (2007)	3.24	1.89	17	0 (0.0)	-	-	-4.205	2.015	-4.205	2.015	-4.205	2.015
Craissati et al. (2008)	2.34	1.97	200	15 (7.5)	0.358	0.114	-3.587	0.502	-2.870	0.336	-1.795	0.317
de Vogel et al. (2004)	6.12	1.7	100	26 (26.0)	0.459	0.150	-3.984	1.029	-3.067	0.740	-1.691	0.346
Eher et al. (2008)	2.30	1.69	151	3 (2.0)	0.695	0.255	-6.456	1.472	-5.065	1.035	-2.979	0.632
Endrass et al. (2009)	3.50	1.67	95	8 (8.4)	0.470	0.205	-4.300	1.023	-3.361	0.661	-1.951	0.391
Epperson (2003)	2.77	2.18	151	16 (10.6)	0.384	0.116	-3.483	0.561	-2.714	0.378	-1.561	0.293
Haag (2005)	3.94	2.01	198	39 (19.7)	0.296	0.094	-2.674	0.472	-2.082	0.307	-1.195	0.188
Hanson et al. (2007)	3.06	2.11	31	0 (0.0)	-	-	-4.812	2.008	-4.812	2.008	-4.812	2.008
Harkins & Beech (2007)	2.78	2.22	198	19 (9.6)	0.377	0.108	-3.576	0.525	-2.823	0.351	-1.692	0.263

Appendix continues.

Appendix B continued

	Static-99 M (SD)		N	N recid (%)	B_1	B_1SE	$B_{0(0)}$	$B_{0(0)}SE$	$B_{0(2)}$	$B_{0(2)}SE$	$B_{0(5)}$	$B_{0(5)}SE$
Hill et al. (2008)	5.01	1.81	73	8 (11.0)	0.172	0.212	-2.991	1.214	-2.648	0.821	-2.134	0.388
Johansen (2007)	2.99	1.97	272	16 (5.9)	0.234	0.120	-3.572	0.532	-3.104	0.342	-2.401	0.294
Knight & Thornton (2007)	4.49	2.24	433	107 (24.7)	0.241	0.054	-2.267	0.296	-1.785	0.201	-1.062	0.114
Långström (2004)	2.43	1.98	1278	69 (5.4)	0.328	0.054	-3.867	0.238	-3.212	0.157	-2.23	0.144
Langton (2003)	3.31	2.09	226	23 (10.2)	0.250	0.106	-3.114	0.495	-2.614	0.320	-1.864	0.240
Milton (2003)	4.81	2.00	93	15 (16.1)	0.428	0.160	-3.932	0.975	-3.076	0.677	-1.793	0.322
Nicholaichuk (2001)	4.54	1.99	168	38 (22.6)	0.391	0.109	-3.162	0.608	-2.38	0.408	-1.206	0.196
Saum (2007)	2.04	1.64	175	55 (31.4)	0.561	0.120	-2.032	0.333	-0.911	0.182	0.771	0.363
Swinburne Romine et al. (2008)	1.88	1.78	570	48 (8.4)	0.279	0.073	-3.023	0.25	-2.464	0.161	-1.627	0.229
Ternowski (2004)	2.14	1.94	247	16 (6.5)	0.246	0.114	-3.307	0.435	-2.814	0.288	-2.076	0.342
Wilson et al. (2007a & b)	5.48	2.09	100	12 (12.0)	0.027	0.148	-2.143	0.881	-2.088	0.612	-2.007	0.319

Note. In samples with no recidivists, B_0 was estimated but B_1 could not be computed.

Appendix C

Static-99 logistic regression analyses at 10 years

	Static-99 *M (SD)*		*N*	*N* recid	B_1	B_1SE	$B_{0(0)}$	$B_{0(0)}SE$	$B_{0(2)}$	$B_{0(2)}SE$	$B_{0(5)}$	$B_{0(5)}SE$
Allan et al. (2007)	2.1	(1.7)	25	5	0.780	0.387	-3.422	1.296	-1.862	0.692	0.479	0.962
Bengtson (2008)	3.8	(2.1)	291	83	0.244	0.067	-1.886	0.307	-1.398	0.196	-0.667	0.144
Boer (2003)	3.2	(2.3)	295	23	0.408	0.100	-4.162	0.548	-3.346	0.374	-2.124	0.226
Brouillette-Alarie & Proulx (2008)	3.8	(2.2)	110	23	0.368	0.115	-2.890	0.585	-2.155	0.390	-1.052	0.253
Craig et al. (2006)	2.4	(1.7)	66	9	0.148	0.201	-2.223	0.652	-1.926	0.387	-1.481	0.585
Craissati et al. (2008)	1.7	(1.7)	66	6	0.190	0.216	-2.680	0.646	-2.299	0.434	-1.729	0.735
de Vogel et al. (2004)	6.3	(1.8)	71	27	0.509	0.170	-3.778	1.153	-2.760	0.825	-1.234	0.378
Epperson (2003)	3.6	(2.5)	36	8	1.068	0.380	-6.521	2.229	-4.384	1.508	-1.179	0.628
Harkins & Beech (2007)	2.9	(2.2)	129	21	0.392	0.114	-3.022	0.531	-2.237	0.347	-1.059	0.275
Hill et al. (2008)	5.1	(1.8)	54	10	0.050	0.201	-1.740	1.095	-1.639	0.726	-1.488	0.352
Johansen (2007)	3.5	(2.2)	62	8	-0.004	0.177	-1.895	0.726	-1.903	0.463	-1.916	0.462
Knight & Thornton (2007)	4.4	(2.3)	353	106	0.206	0.054	-1.791	0.285	-1.378	0.192	-0.758	0.119
Långström (2004)	2.5	(2.0)	353	26	0.484	0.091	-4.208	0.452	-3.240	0.305	-1.787	0.227
Langton (2003)	3.1	(2.1)	47	6	0.418	0.225	-3.486	1.084	-2.650	0.702	-1.397	0.481
Milton (2003)	5.0	(2.0)	68	17	0.404	0.161	-3.286	0.965	-2.478	0.667	-1.267	0.318
Nicholaichuk (2001)	4.7	(2.1)	59	15	0.274	0.160	-2.428	0.884	-1.880	0.595	-1.058	0.309
Swinburne Romine et al. (2008)	1.9	(1.8)	543	61	0.232	0.069	-2.568	0.218	-2.105	0.141	-1.410	0.223
Wilson et al. (2007a & b)	5.6	(2.5)	15	1	17.40	3473.29	-156.64	31259.6	-121.83	24313.0	-69.62	13893

Appendix D

Logistic Regression Analyses of the Static-99 centered on 2 for Five-year Sexual Recidivism

Table D-1. *Recidivism Criteria*

	Static-99 *M (SD)*		*N*	*N* recid (%)	B_1	$B_1 SE$	B_0	B_0 95% C.I.	
Charges									
Allan et al. (2007)	2.3	2.0	299	35 (11.7)	0.411	0.086	8.3	5.5	12.5
Bartosh et al. (2003)	3.2	2.2	90	12 (13.3)	0.138	0.135	11.2	5.5	21.2
Bengtson (2008)	3.8	2.1	310	61 (19.7)	0.210	0.070	13.8	9.5	19.7
Bigras (2007)	2.5	1.9	207	19 (9.2)	0.362	0.117	6.4	3.6	11.2
Cortoni & Nunes (2007)	3.2	1.9	17	0 (0.0)	-	-	1.8	<0.1	43.6
Epperson (2003)	2.8	2.2	151	16 (10.6)	0.384	0.116	6.2	3.1	12.2
Hanson et al. (2007)	3.1	2.1	31	0 (0.0)	-	-	0.8	<0.1	29.4
Johansen (2007)	3.0	2.0	272	16 (5.9)	0.234	0.120	4.3	2.2	8.1
Knight & Thornton (2007)	4.5	2.2	433	107 (24.7)	0.241	0.054	14.4	10.2	19.9
Ternowski (2004)	2.1	1.9	247	16 (6.5)	0.246	0.114	5.7	3.3	9.5
Wilson et al. (2007a & b)	5.5	2.1	100	12 (12.0)	0.027	0.148	11.0	3.6	29.1
Convictions									
Boer (2003)	3.3	2.3	299	11 (3.7)	0.515	0.146	1.0	0.3	3.4
Bonta & Yessine (2005)	5.2	2.0	101	19 (18.8)	0.265	0.139	8.4	2.8	22.5
Brouillette-Alarie & Proulx (2008)	3.8	2.2	199	29 (14.6)	0.360	0.095	6.7	3.5	12.5
Craissati et al. (2008)	2.3	2.0	200	15 (7.5)	0.358	0.114	5.4	2.8	9.9
de Vogel et al. (2004)	6.1	1.7	100	26 (26.0)	0.459	0.150	4.4	1.1	16.6
Eher et al. (2008)	2.3	1.7	151	3 (2.0)	0.695	0.255	0.6	0.1	4.6
Endrass et al. (in press)	3.5	1.7	95	8 (8.4)	0.470	0.205	3.4	0.9	11.3
Haag (2005)	3.9	2.0	198	39 (19.7)	0.296	0.094	11.1	6.4	18.5
Harkins & Beech (2007)	2.8	2.2	198	19 (9.6)	0.377	0.108	5.6	2.9	10.6
Hill et al. (2008)	5.0	1.8	73	8 (11.0)	0.172	0.212	6.6	1.4	26.1
Långström (2004)	2.4	2.0	1278	69 (5.4)	0.328	0.054	3.9	2.9	5.2
Langton (2003)	3.3	2.1	226	23 (10.2)	0.250	0.106	6.8	3.8	12.1
Milton (2003)	4.8	2.0	93	15 (16.1)	0.428	0.160	4.4	1.2	14.8
Nicholaichuk (2001)	4.5	2.0	168	38 (22.6)	0.391	0.109	8.5	4.0	17.1
Swinburne Romine et al. (2008)	1.9	1.8	570	48 (8.4)	0.279	0.073	7.8	5.8	10.4

Table D-2. *Number of Recidivism Sources*

	Static-99 *M*	(*SD*)	*N*	*N* recid (%)	B_1	B_1SE	B_0	B_0 95% C.I.	
One									
Allan et al. (2007)	2.3	2.0	299	35 (11.7)	0.411	0.086	8.3	5.5	12.5
Bartosh et al. (2003)	3.2	2.2	90	12 (13.3)	0.138	0.135	11.2	5.5	21.2
Bengtson (2008)	3.8	2.1	310	61 (19.7)	0.210	0.070	13.8	9.5	19.7
Bigras (2007)	2.5	1.9	207	19 (9.2)	0.362	0.117	6.4	3.6	11.2
Boer (2003)	3.3	2.3	299	11 (3.7)	0.515	0.146	1.0	0.3	3.4
Cortoni & Nunes (2007)	3.2	1.9	17	0 (0.0)	-	-	1.8	0.03	43.6
de Vogel et al. (2004)	6.1	1.7	100	26 (26.0)	0.459	0.150	4.4	1.1	16.6
Eher et al. (2008)	2.3	1.7	151	3 (2.0)	0.695	0.255	0.6	0.1	4.6
Endrass et al. (in press)	3.5	1.7	95	8 (8.4)	0.470	0.205	3.4	0.9	11.3
Epperson (2003)	2.8	2.2	151	16 (10.6)	0.384	0.116	6.2	3.1	12.2
Haag (2005)	3.9	2.0	198	39 (19.7)	0.296	0.094	11.1	6.4	18.5
Hill et al. (2008)	5.0	1.8	73	8 (11.0)	0.172	0.212	6.6	1.4	26.1
Långström (2004)	2.4	2.0	1278	69 (5.4)	0.328	0.054	3.9	2.9	5.2
Langton (2003)	3.3	2.1	226	23 (10.2)	0.250	0.106	6.8	3.8	12.1
Nicholaichuk (2001)	4.5	2.0	168	38 (22.6)	0.391	0.109	8.5	4.0	17.1
Swinburne Romine et al. (2008)	1.9	1.8	570	48 (8.4)	0.279	0.073	7.8	5.8	10.4
Ternowski (2004)	2.1	1.9	247	16 (6.5)	0.246	0.114	5.7	3.3	9.5
Wilson et al. (2007a & b)	5.5	2.1	100	12 (12.0)	0.027	0.148	11.0	3.6	29.1
Two									
Bonta & Yessine (2005)	5.2	2.0	101	19 (18.8)	0.265	0.139	8.4	2.8	22.5
Brouillette-Alarie & Proulx (2008)	3.8	2.2	199	29 (14.6)	0.360	0.095	6.7	3.5	12.5
Harkins & Beech (2007)	2.8	2.2	198	19 (9.6)	0.377	0.108	5.6	2.9	10.6
Johansen (2007)	3.0	2.0	272	16 (5.9)	0.234	0.120	4.3	2.2	8.1
Milton (2003)	4.8	2.0	93	15 (16.1)	0.428	0.160	4.4	1.2	14.8
Four									
Craissati et al. (2008)	2.3	2.0	200	15 (7.5)	0.358	0.114	5.4	2.8	9.9
Hanson et al. (2007)	3.1	2.1	31	0 (0.0)	-	-	0.8	0.02	29.4
Knight & Thornton (2007)	4.5	2.2	433	107 (24.7)	0.241	0.054	14.4	10.2	19.9

Table D-3. *Number of Recidivism Sources (dichotomized)*

	Static-99 *M (SD)*		*N*	*N* recid (%)	B_1	$B_1 SE$	B_0	B_0 95% C.I.	
One									
Allan et al. (2007)	2.3	2.0	299	35 (11.7)	0.411	0.086	8.3	5.5	12.5
Bartosh et al. (2003)	3.2	2.2	90	12 (13.3)	0.138	0.135	11.2	5.5	21.2
Bengtson (2008)	3.8	2.1	310	61 (19.7)	0.210	0.070	13.8	9.5	19.7
Bigras (2007)	2.5	1.9	207	19 (9.2)	0.362	0.117	6.4	3.6	11.2
Boer (2003)	3.3	2.3	299	11 (3.7)	0.515	0.146	1.0	0.3	3.4
Cortoni & Nunes (2007)	3.2	1.9	17	0 (0.0)	-	-	1.8	0.03	43.6
de Vogel et al. (2004)	6.1	1.7	100	26 (26.0)	0.459	0.150	4.4	1.1	16.6
Eher et al. (2008)	2.3	1.7	151	3 (2.0)	0.695	0.255	0.6	0.1	4.6
Endrass et al. (in press)	3.5	1.7	95	8 (8.4)	0.470	0.205	3.4	0.9	11.3
Epperson (2003)	2.8	2.2	151	16 (10.6)	0.384	0.116	6.2	3.1	12.2
Haag (2005)	3.9	2.0	198	39 (19.7)	0.296	0.094	11.1	6.4	18.5
Hill et al. (2008)	5.0	1.8	73	8 (11.0)	0.172	0.212	6.6	1.4	26.1
Långström (2004)	2.4	2.0	1278	69 (5.4)	0.328	0.054	3.9	2.9	5.2
Langton (2003)	3.3	2.1	226	23 (10.2)	0.250	0.106	6.8	3.8	12.1
Nicholaichuk (2001)	4.5	2.0	168	38 (22.6)	0.391	0.109	8.5	4.0	17.1
Swinburne Romine et al. (2008)	1.9	1.8	570	48 (8.4)	0.279	0.073	7.8	5.8	10.4
Ternowski (2004)	2.1	1.9	247	16 (6.5)	0.246	0.114	5.7	3.3	9.5
Wilson et al. (2007a & b)	5.5	2.1	100	12 (12.0)	0.027	0.148	11.0	3.6	29.1
Two or more									
Bonta & Yessine (2005)	5.2	2.0	101	19 (18.8)	0.265	0.139	8.4	2.8	22.5
Brouillette-Alarie & Proulx (2008)	3.8	2.2	199	29 (14.6)	0.360	0.095	6.7	3.5	12.5
Craissati et al. (2008)	2.3	2.0	200	15 (7.5)	0.358	0.114	5.4	2.8	9.9
Hanson et al. (2007)	3.1	2.1	31	0 (0.0)	-	-	0.8	0.02	29.4
Harkins & Beech (2007)	2.8	2.2	198	19 (9.6)	0.377	0.108	5.6	2.9	10.6
Johansen (2007)	3.0	2.0	272	16 (5.9)	0.234	0.120	4.3	2.2	8.1
Knight & Thornton (2007)	4.5	2.2	433	107 (24.7)	0.241	0.054	14.4	10.2	19.9
Milton (2003)	4.8	2.0	93	15 (16.1)	0.428	0.160	4.4	1.2	14.8

Table D-4. *Use of national records to measure recidivism*

	Static-99 *M (SD)*		*N*	*N* recid (%)	B_1	B_1SE	B_0	B_0 *95% C.I.*	
No									
Epperson (2003)	2.8	2.2	151	16 (10.6)	0.384	0.116	6.2	3.1	12.2
Johansen (2007)	3.0	2.0	272	16 (5.9)	0.234	0.120	4.3	2.2	8.1
Swinburne Romine et al. (2008)	1.9	1.8	570	48 (8.4)	0.279	0.073	7.8	5.8	10.4
Ternowski (2004)	2.1	1.9	247	16 (6.5)	0.246	0.114	5.7	3.3	9.5
Yes									
Allan et al. (2007)	2.3	2.0	299	35 (11.7)	0.411	0.086	8.3	5.5	12.5
Bartosh et al. (2003)	3.2	2.2	90	12 (13.3)	0.138	0.135	11.2	5.5	21.2
Bengtson (2008)	3.8	2.1	310	61 (19.7)	0.210	0.070	13.8	9.5	19.7
Bigras (2007)	2.5	1.9	207	19 (9.2)	0.362	0.117	6.4	3.6	11.2
Boer (2003)	3.3	2.3	299	11 (3.7)	0.515	0.146	1.0	0.3	3.4
Bonta & Yessine (2005)	5.2	2.0	101	19 (18.8)	0.265	0.139	8.4	2.8	22.5
Brouillette-Alarie & Proulx (2008)	3.8	2.2	199	29 (14.6)	0.360	0.095	6.7	3.5	12.5
Cortoni & Nunes (2007)	3.2	1.9	17	0 (0.0)	-	-	1.8	<0.1	43.6
Craissati et al. (2008)	2.3	2.0	200	15 (7.5)	0.358	0.114	5.4	2.8	9.9
de Vogel et al. (2004)	6.1	1.7	100	26 (26.0)	0.459	0.150	4.4	1.1	16.6
Eher et al. (2008)	2.3	1.7	151	3 (2.0)	0.695	0.255	0.6	0.1	4.6
Endrass et al. (in press)	3.5	1.7	95	8 (8.4)	0.470	0.205	3.4	0.9	11.3
Haag (2005)	3.9	2.0	198	39 (19.7)	0.296	0.094	11.1	6.4	18.5
Hanson et al. (2007)	3.1	2.1	31	0 (0.0)	-	-	0.8	<0.1	29.4
Harkins & Beech (2007)	2.8	2.2	198	19 (9.6)	0.377	0.108	5.6	2.9	10.6
Hill et al. (2008)	5.0	1.8	73	8 (11.0)	0.172	0.212	6.6	1.4	26.1
Knight & Thornton (2007)	4.5	2.2	433	107 (24.7)	0.241	0.054	14.4	10.2	19.9
Långström (2004)	2.4	2.0	1278	69 (5.4)	0.328	0.054	3.9	2.9	5.2
Langton (2003)	3.3	2.1	226	23 (10.2)	0.250	0.106	6.8	3.8	12.1
Milton (2003)	4.8	2.0	93	15 (16.1)	0.428	0.160	4.4	1.2	14.8
Nicholaichuk (2001)	4.5	2.0	168	38 (22.6)	0.391	0.109	8.5	4.0	17.1
Wilson et al. (2007a & b)	5.5	2.1	100	12 (12.0)	0.027	0.148	11.0	3.6	29.1

Table D-5. *Was street time used?*

	Static-99 *M*	*(SD)*	*N*	*N* recid (%)	B_1	B_1SE	B_0	B_0 95% C.I.	
No									
Allan et al. (2007)	2.2	2.0	299	35 (11.7)	0.411	0.086	8.3	5.5	12.5
Bartosh et al. (2003)	3.2	2.2	90	12 (13.3)	0.138	0.135	11.2	5.5	21.2
Bengtson (2008)	3.8	2.1	310	61 (19.7)	0.210	0.070	13.8	9.5	19.7
Bigras (2007)	2.5	1.9	207	19 (9.2)	0.362	0.117	6.4	3.6	11.2
Boer (2003)	3.3	2.3	299	11 (3.7)	0.515	0.146	1.0	0.3	3.4
Bonta & Yessine (2005)	5.2	2.0	101	19 (18.8)	0.265	0.139	8.4	2.8	22.5
Brouillette-Alarie & Proulx (2008)	3.8	2.2	199	29 (14.6)	0.360	0.095	6.7	3.5	12.5
Cortoni & Nunes (2007)	3.2	1.9	17	0 (0.0)	-	-	1.8	<0.1	43.6
Craissati et al. (2008)	2.3	2.0	200	15 (7.5)	0.358	0.114	5.4	2.8	9.9
de Vogel et al. (2004)	6.1	1.7	100	26 (26.0)	0.459	0.150	4.4	1.1	16.6
Eher et al. (2008)	2.3	1.7	151	3 (2.0)	0.695	0.255	0.6	0.1	4.6
Endrass et al. (in press)	3.5	1.7	95	8 (8.4)	0.470	0.205	3.4	0.9	11.3
Epperson (2003)	2.8	2.2	151	16 (10.6)	0.384	0.116	6.2	3.1	12.2
Haag (2005)	3.9	2.0	198	39 (19.7)	0.296	0.094	11.1	6.4	18.5
Harkins & Beech (2007)	2.8	2.2	198	19 (9.6)	0.377	0.108	5.6	2.9	10.6
Johansen (2007)	3.0	2.0	272	16 (5.9)	0.234	0.120	4.3	2.2	8.1
Långström (2004)	2.4	2.0	1278	69 (5.4)	0.328	0.054	3.9	2.9	5.2
Langton (2003)	3.3	2.1	226	23 (10.2)	0.250	0.106	6.8	3.8	12.1
Milton (2003)	4.8	2.0	93	15 (16.1)	0.428	0.160	4.4	1.2	14.8
Nicholaichuk (2001)	4.5	2.0	168	38 (22.6)	0.391	0.109	8.5	4.0	17.1
Swinburne Romine et al. (2008)	1.9	1.8	570	48 (8.4)	0.279	0.073	7.8	5.8	10.4
Ternowski (2004)	2.1	1.9	247	16 (6.5)	0.246	0.114	5.7	3.3	9.5
Wilson et al. (2007a & b)	5.5	2.1	100	12 (12.0)	0.027	0.148	11.0	3.6	29.1
Yes									
Hanson et al. (2007)	3.1	2.1	31	0 (0.0)	-	-	0.8	<0.1	29.4
Hill et al. (2008)	5.0	1.8	73	8 (11.0)	0.172	0.212	6.6	1.4	26.1
Knight & Thornton (2007)	4.5	2.2	433	107 (24.7)	0.241	0.054	14.4	10.2	19.9

Table D-6. *Did they cite the coding rules?*

	Static-99 *M (SD)*		*N*	*N* recid (%)	B_1	$B_1 SE$	B_0	B_0 95% C.I.	
No									
Bigras (2007)	2.5	1.9	207	19 (9.2)	0.362	0.117	6.4	3.6	11.2
Craissati et al. (2008)	2.3	2.0	200	15 (7.5)	0.358	0.114	5.4	2.8	9.9
de Vogel et al. (2004)	6.1	1.7	100	26 (26.0)	0.459	0.150	4.4	1.1	16.6
Epperson (2003)	2.8	2.2	151	16 (10.6)	0.384	0.116	6.2	3.1	12.2
Harkins & Beech (2007)	2.8	2.2	198	19 (9.6)	0.377	0.108	5.6	2.9	10.6
Hill et al. (2008)	5.0	1.8	73	8 (11.0)	0.172	0.212	6.6	1.4	26.1
Langton (2003)	3.3	2.1	226	23 (10.2)	0.250	0.106	6.8	3.8	12.1
Milton (2003)	4.8	2.0	93	15 (16.1)	0.428	0.160	4.4	1.2	14.8
Nicholaichuk (2001)	4.5	2.0	168	38 (22.6)	0.391	0.109	8.5	4.0	17.1
Wilson et al. (2007a & b)	5.5	2.1	100	12 (12.0)	0.027	0.148	11.0	3.6	29.1
Yes									
Allan et al. (2007)	2.3	2.0	299	35 (11.7)	0.411	0.086	8.3	5.5	12.5
Bartosh et al. (2003)	3.2	2.2	90	12 (13.3)	0.138	0.135	11.2	5.5	21.2
Bengtson (2008)	3.8	2.1	310	61 (19.7)	0.210	0.070	13.8	9.5	19.7
Boer (2003)	3.3	2.2	299	11 (3.7)	0.515	0.146	1.0	0.3	3.4
Bonta & Yessine (2005)	5.2	2.0	101	19 (18.8)	0.265	0.139	8.4	2.8	22.5
Brouillette-Alarie & Proulx (2008)	3.8	2.2	199	29 (14.6)	0.360	0.095	6.7	3.5	12.5
Cortoni & Nunes (2007)	3.2	1.9	17	0 (0.0)	-	-	1.8	<0.1	43.6
Eher et al. (2008)	2.3	1.7	151	3 (2.0)	0.695	0.255	0.6	0.1	4.6
Endrass et al. (in press)	3.5	1.7	95	8 (8.4)	0.470	0.205	3.4	0.9	11.3
Haag (2005)	3.9	2.0	198	39 (19.7)	0.296	0.094	11.1	6.4	18.5
Hanson et al. (2007)	3.1	2.1	31	0 (0.0)	-	-	0.8	<0.1	29.4
Johansen (2007)	3.0	2.0	272	16 (5.9)	0.234	0.120	4.3	2.2	8.1
Knight & Thornton (2007)	4.5	2.2	433	107 (24.7)	0.241	0.054	14.4	10.2	19.9
Långström (2004)	2.4	2.0	1278	69 (5.4)	0.328	0.054	3.9	2.9	5.2
Swinburne Romine et al. (2008)	1.9	1.8	570	48 (8.4)	0.279	0.073	7.8	5.8	10.4
Ternowski (2004)	2.1	1.9	247	16 (6.5)	0.246	0.114	5.7	3.3	9.5

Table D-7. *Federal vs Provincial*

	Static-99 *M*	*(SD)*	*N*	*N* recid (%)	B_1	B_1SE	B_0	B_0 95% C.I.	
Provincial									
*Bonta & Yessine (2005)	5.4	2.0	41	8 (19.5)	0.403	0.238	4.9	0.6	30.2
*Brouillette-Alarie & Proulx (2008)	3.3	2.1	105	14 (13.3)	0.358	0.141	7.4	3.3	16.0
*Hanson et al. (2007)	2.7	1.9	25	0 (0.0)	-	-	1.0	<0.1	34.2
Ternowski (2004)	2.1	1.9	247	16 (6.5)	0.246	0.114	5.7	3.3	9.5
Federal									
*Bigras (2007)	2.5	1.9	207	19 (9.2)	0.362	0.117	6.4	3.6	11.2
*Boer (2003)	3.3	2.3	299	11 (3.7)	0.515	0.146	1.0	0.3	3.4
Bonta & Yessine (2005)	4.9	2.0	54	9 (16.7)	0.165	0.188	10.7	2.8	33.1
Brouillette-Alarie & Proulx (2008)	4.4	2.2	94	15 (16.0)	0.385	0.138	5.7	1.9	15.5
*Cortoni & Nunes (2007)	3.2	1.9	17	0 (0.0)	-	-	1.5	<0.1	43.6
*Haag (2005)	3.9	2.0	198	39 (19.7)	0.296	0.094	11.1	6.4	18.5
*Langton (2003)	3.3	2.1	226	23 (10.2)	0.250	0.106	6.8	3.8	12.1
*Nicholaichuk (2001)	4.5	2.0	168	38 (22.6)	0.391	0.109	8.5	4.0	17.1
*Wilson et al. (2007a, b)	5.5	2.1	100	12 (12.0)	0.027	0.148	11.0	3.6	29.1

* Refers to a subsample of the full study.

Table D-8. *Offender Type*

	Static-99 M (SD)		N	N recid (%)	B_1	B_1SE	B_0	B_0 95% C.I.	
Child molester									
Allan et al. (2007)	2.3	2.0	299	35 (11.7)	0.411	0.086	8.3	5.5	12.5
*Bigras (2007)	2.2	2.0	90	7 (7.8)	0.306	0.177	6.4	2.7	14.4
*Boer (2003)	2.0	1.9	164	2 (1.2)	0.408	0.287	0.9	0.2	4.9
*Brouillette-Alarie & Proulx (2008)	3.5	2.3	117	19 (16.2)	0.448	0.126	6.7	2.9	14.6
*Craissati et al. (2008)	3.0	1.6	48	4 (8.3)	0.387	0.309	5.0	1.1	19.3
*Eher et al. (2008)	1.8	1.4	85	1 (1.2)	28.115	1612.3	0.0	<0.1	>99
*Haag (2005)	3.5	2.2	71	13 (18.3)	0.188	0.140	13.9	6.6	26.9
*Hanson et al. (2007)	2.3	1.7	16	0 (0.0)	-	-	1.6	<0.1	45.2
*Knight & Thornton (2007)	4.1	2.5	211	46 (21.8)	0.206	0.070	14.3	9.1	21.8
*Långström (2004)	1.6	1.7	580	25 (4.3)	0.437	0.088	3.8	2.4	5.8
*Nicholaichuk (2001)	3.6	2.1	48	11 (22.9)	0.487	0.192	9.7	3.1	26.3
*Swinburne Romine et al. (2008)	1.5	1.7	332	32 (9.6)	0.214	0.096	10.1	7.2	13.9
*Wilson et al. (2007a, b)	5.5	2.4	38	5 (13.2)	0.268	0.238	5.0	0.5	34.0
Rapist									
*Bigras (2007)	2.9	1.8	81	8 (9.9)	0.513	0.202	4.6	1.5	13.3
*Boer (2003)	4.8	1.7	123	8 (6.5)	0.519	0.213	1.2	0.2	6.9
*Brouillette-Alarie & Proulx (2008)	4.1	2.1	64	8 (12.5)	0.271	0.168	6.7	2.0	19.9
*Craissati et al. (2008)	1.7	1.7	120	5 (4.2)	0.144	0.229	4.2	1.8	9.7
*Eher et al. (2008)	2.8	1.5	60	0 (0.0)	-	-	0.4	<0.1	17.5
*Haag (2005)	4.1	1.8	118	25 (21.2)	0.393	0.139	9.3	4.1	19.8
*Hanson et al. (2007)	3.6	2.2	11	0 (0.0)	-	-	2.3	<0.1	55.1
*Knight & Thornton (2007)	4.8	1.7	148	33 (22.3)	0.244	0.119	12.1	5.6	24.2
*Långström (2004)	2.8	1.7	544	34 (6.2)	0.318	0.091	4.3	2.8	6.7
*Nicholaichuk (2001)	4.8	1.8	92	19 (20.6)	0.374	0.161	7.3	2.3	21.2
*Swinburne Romine et al. (2008)	2.7	1.4	62	5 (8.1)	0.740	0.360	3.4	0.8	14.1
*Wilson et al. (2007a, b)	5.5	1.9	37	5 (13.5)	-0.116	0.270	18.8	3.2	62.2

* Refers to a subsample of the full study.

Table D-9. *Country*

	Static-99 *M*	*(SD)*	*N*	*N* recid (%)	B_1	B_1SE	B_0	B_0 95% C.I.	
Canada									
Bigras (2007)	2.5	1.9	207	19 (9.2)	0.362	0.117	6.4	3.6	11.2
Boer (2003)	3.3	2.3	299	11 (3.7)	0.515	0.146	1.0	0.3	3.4
Bonta & Yessine (2005)	5.2	2.0	101	19 (18.8)	0.265	0.139	8.4	2.8	22.5
Brouillette-Alarie & Proulx (2008)	3.8	2.2	199	29 (14.6)	0.360	0.095	6.7	3.5	12.5
Cortoni & Nunes (2007)	3.2	1.9	17	0 (0.0)	-	-	1.8	<0.1	43.6
Haag (2005)	3.9	2.0	198	39 (19.7)	0.296	0.094	11.1	6.4	18.5
Hanson et al. (2007)	3.1	2.1	31	0 (0.0)	-	-	0.8	<0.1	29.4
Langton (2003)	3.3	2.1	226	23 (10.2)	0.250	0.106	6.8	3.8	12.1
Nicholaichuk (2001)	4.5	2.0	168	38 (22.6)	0.391	0.109	8.5	4.0	17.1
Ternowski (2004)	2.1	1.9	247	16 (6.5)	0.246	0.114	5.7	3.3	9.5
Wilson et al. (2007a & b)	5.5	2.0	100	12 (12.0)	0.027	0.148	11.0	3.6	29.1
United States									
Bartosh et al. (2003)	3.2	2.2	90	12 (13.3)	0.138	0.135	11.2	5.5	21.2
Epperson (2003)	2.8	2.2	151	16 (10.6)	0.384	0.116	6.2	3.1	12.2
Johansen (2007)	3.0	2.0	272	16 (5.9)	0.234	0.120	4.3	2.2	8.1
Knight/Thornton (2007)	4.5	2.2	433	107 (24.7)	0.241	0.054	14.4	10.2	19.9
Swinburne Romine (2008	1.9	1.8	570	48 (8.4)	0.279	0.073	7.8	5.8	10.4
United Kingdom									
Craissati et al. (2008)	2.3	2.0	200	15 (7.5)	0.358	0.114	5.4	2.8	9.9
Harkins & Beech (2007)	2.8	2.2	198	19 (9.6)	0.377	0.108	5.6	2.9	10.6
Milton (2003)	4.8	2.0	93	15 (16.1)	0.428	0.160	4.4	1.2	14.8
Continental Europe									
Bengtson (2008)	3.8	2.1	310	61 (19.7)	0.210	0.070	13.8	9.5	19.7
de Vogel et al. (2004)	6.1	1.7	100	26 (26.0)	0.459	0.150	4.4	1.1	16.6
Eher et al. (2008)	2.3	1.7	151	3 (2.0)	0.695	0.255	0.6	0.1	4.6
Endrass et al. (in press)	3.5	1.7	95	8 (8.4)	0.470	0.205	3.4	0.9	11.3
Hill et al. (2008)	5.0	1.8	73	8 (11.0)	0.172	0.212	6.6	1.4	26.1
Långström (2004)	2.4	2.0	1278	69 (5.4)	0.328	0.054	3.9	2.9	5.2
New Zealand									
Allan et al. (2007)	2.3	2.0	299	35 (11.7)	0.411	0.086	8.3	5.5	12.5

Table D-10. *Aboriginal*

	Static-99 *M*	(*SD*)	*N*	*N* recid (%)	B_1	$B_1 SE$	B_0	B_0 95% C.I.	
Non-Aboriginal									
*Boer (2003)	3.3	2.3	228	11 (4.8)	0.486	0.140	1.5	0.5	4.7
*Bonta & Yessine (2005)	5.4	2.0	83	15 (18.1)	0.282	0.159	7.2	1.9	23.6
*Brouillette-Alarie & Proulx (2008)	3.8	2.2	196	27 (13.8)	0.359	0.097	6.3	3.2	12.0
*Haag (2005)	4.0	2.1	146	25 (17.1)	0.316	0.112	8.7	4.3	17.1
*Hanson et al. (2007)	3.3	2.3	23	0 (0.0)	-	-	1.1	0.0	36.1
Hill et al. (2008)	5.0	1.8	73	8 (11.0)	0.172	0.212	6.6	1.4	26.1
*Nicholaichuk (2001)	4.6	2.1	127	27 (21.2)	0.368	0.124	8.2	3.4	18.5
*Swinburne Romine et al. (2008)	1.9	1.8	535	42 (7.8)	0.286	0.077	7.3	5.3	9.9
Aboriginal									
*Boer (2003)	3.3	2.2	56	0 (0.0)	-	-	0.4	0.0	18.6
*Bonta & Yessine (2005)	4.5	1.8	18	4 (22.2)	0.311	0.327	10.8	1.3	53.0
*Brouillette-Alarie & Proulx (2008)	5.0	0.0	3	2 (66.7)	-	-	66.7	15.4	95.7
*Haag (2005)	3.9	1.6	50	14 (28.0)	0.300	0.204	17.4	6.7	38.3
*Hanson et al. (2007)	2.3	1.5	6	0 (0.0)	-	-	4.2	0.1	70.4
*Nicholaichuk (2001)	4.4	1.8	41	11 (26.8)	0.519	0.243	8.0	1.6	31.2
*Swinburne Romine et al. (2008)	1.7	2.1	3	0 (0.0)	-	-	11.1	0.3	82.1

* Refers to a subsample of the full study.

Table D-11. *NonWhite*

	Static-99 *M (SD)*		*N*	*N* recid (%)	B_1	B_1SE	B_0	B_0 95% C.I.	
White									
*Boer (2003)	3.2	2.3	228	11 (4.8)	0.486	0.140	1.5	0.5	4.7
*Bonta & Yessine (2005)	5.4	2.0	77	14 (18.2)	0.358	0.176	5.4	1.2	21.3
*Brouillette-Alarie & Proulx (2008)	3.7	2.2	192	26 (13.5)	0.354	0.098	6.4	3.2	12.1
*Haag (2005)	4.0	2.1	141	25 (17.7)	0.303	0.112	9.3	4.5	18.0
Hill et al. (2008)	5.0	1.8	73	8 (11.0)	0.172	0.212	6.6	1.4	26.1
*Knight & Thornton (2007)	4.4	2.2	400	93 (23.2)	0.246	0.058	13.5	9.3	19.0
*Swinburne Romine et al. (2008)	1.8	1.8	507	40 (7.9)	0.292	0.080	7.4	5.3	10.1
NonWhite									
*Boer (2003)	3.3	2.2	56	0 (0.0)	-	-	0.4	<0.1	18.6
*Bonta & Yessine (2005)	4.7	2.0	24	5 (20.8)	0.104	0.250	16.4	3.4	52.6
*Brouillette-Alarie & Proulx (2008)	5.1	2.2	7	3 (42.8)	0.241	0.329	25.6	1.4	88.9
*Haag (2005)	3.8	1.7	55	14 (25.4)	0.345	0.200	14.8	5.7	33.2
*Hanson et al. (2007)	2.3	1.5	6	0 (0.0)	-	-	4.2	0.1	70.4
*Knight & Thornton (2007)	5.6	2.0	33	14 (42.4)	0.064	0.183	36.9	11.9	71.7
*Swinburne Romine et al. (2008)	2.3	2.1	31	2 (6.4)	0.264	0.324	5.2	1.0	23.1

* Refers to a subsample of the full study.

Table D-12. *Non-Aboriginal non-white*

	Static-99 *M (SD)*		*N*	*N* recid (%)	B_1	B_1SE	B_0	B_0 95% C.I.	
White									
*Boer (2003)	3.2	2.3	228	11 (4.8)	0.486	0.140	1.5	0.5	4.7
*Bonta & Yessine (2005)	5.4	2.0	77	14 (18.2)	0.358	0.176	5.4	1.2	21.3
*Brouillette-Alarie & Proulx (2008)	3.7	2.2	192	26 (13.5)	0.354	0.098	6.4	3.2	12.1
*Haag (2005)	4.0	2.1	141	25 (17.7)	0.303	0.112	9.3	4.5	18.0
Hill et al. (2008)	5.0	1.8	73	8 (11.0)	0.172	0.212	6.6	1.4	26.1
*Knight & Thornton (2007)	4.4	2.2	400	93 (23.3)	0.246	0.058	13.4	9.3	19.0
*Swinburne Romine et al. (2008)	1.8	1.8	507	40 (7.9)	0.292	0.080	7.4	5.3	10.1
NonAboriginal NonWhite									
*Bonta & Yessine (2005)	5.2	2.8	6	1 (16.7)	-0.222	0.454	26.7	1.8	87.7
*Brouillette-Alarie & Proulx (2008)	3.7	2.2	4	1 (25.0)	0.508	0.789	4.0	0.0	99.0
*Haag (2005)	2.4	1.8	5	0 (0.0)			5.0	0.1	74.6
*Knight & Thornton (2007)	5.6	2.0	33	14 (42.4)	0.064	0.183	36.9	11.9	71.7
*Swinburne Romine et al. (2008)	2.4	2.1	28	2 (7.1)	0.247	0.324	5.8	1.1	25.3

* Refers to a subsample of the full study.

Table D-13. *Treated*

	Static-99 *M* (*SD*)		*N*	*N* recid (%)	B_1	$B_1 SE$	B_0	B_0 95% C.I.	
Didn't Start									
*Craissati et al. (2008)	2.2	1.6	66	7 (10.6)	0.400	0.244	8.6	3.6	19.2
*Haag (2005)	3.8	1.9	73	13 (17.8)	0.135	0.159	14.2	6.3	29.0
*Swinburne Romine et al. (2008)	1.5	1.6	128	8 (6.3)	0.173	0.197	6.5	3.3	12.5
*Ternowski (2004)	1.8	1.5	45	4 (8.9)	0.212	0.323	8.9	3.4	21.7
Started but dropped out									
*Haag (2005)	4.6	2.0	22	6 (27.3)	0.890	0.413	2.2	0.1	33.6
*Swinburne Romine et al. (2008)	2.1	1.9	289	34 (11.8)	0.356	0.091	9.7	6.6	14.0
*Ternowski (2004)	3.8	2.6	18	3 (16.7)	0.161	0.245	12.4	2.5	44.4
Completed									
*Allan et al. (2007)	2.3	2.0	299	35 (11.7)	0.411	0.086	8.3	5.5	12.5
*Craissati et al. (2008)	2.4	2.2	134	8 (6.0)	0.402	0.141	3.6	1.4	9.0
*Haag (2005)	3.9	2.1	103	20 (19.4)	0.288	0.129	11.2	5.2	22.3
*Swinburne Romine et al. (2008)	1.7	1.6	145	6 (4.1)	-0.364	0.344	3.3	1.2	8.8
*Ternowski (2004)	2.1	1.9	184	9 (4.9)	0.252	0.153	4.3	2.1	8.6

* Refers to a subsample of the full study.

Table D-14. *Started Treatment*

	Static-99 *M*	*(SD)*	*N*	*N* recid (%)	B_1	B_1SE	B_0	B_0 95% C.I.	
No									
*Bonta & Yessine (2005)	4.8	1.9	48	11 (22.9)	0.154	0.186	15.8	4.7	41.5
*Craissati et al. (2008)	2.2	1.6	66	7 (10.6)	0.400	0.244	8.6	3.6	19.2
*Haag (2005)	3.8	1.9	73	13 (17.8)	0.135	0.159	14.2	6.3	29.0
*Swinburne Romine et al. (2008)	1.5	1.6	128	8 (6.3)	0.173	0.197	6.5	3.3	12.5
*Ternowski (2004)	1.8	1.5	45	4 (8.9)	0.212	0.323	8.9	3.4	21.7
Yes									
Allan et al. (2007)	2.3	2.0	299	35 (11.7)	0.411	0.086	8.3	5.5	12.5
*Bonta & Yessine (2005)	5.5	2.0	53	8 (15.1)	0.526	0.246	1.9	0.2	18.0
Cortoni & Nunes (2007)	3.2	1.9	17	0 (0.0)	-	-	1.5	0.0	43.6
*Craissati et al. (2008)	2.4	2.2	134	8 (6.0)	0.402	0.141	3.6	1.4	9.0
de Vogel et al. (2004)	6.1	1.7	100	26 (26.0)	0.459	0.150	4.4	1.1	16.6
*Haag (2005)	4.0	2.1	125	26 (20.8)	0.378	0.120	9.4	4.4	18.9
Harkins & Beech (2007)	2.8	2.2	198	19 (9.6)	0.377	0.108	5.6	2.9	10.6
Johansen (2007)	3.0	2.0	272	16 (5.9)	0.234	0.120	4.3	2.2	8.1
*Knight & Thornton (2007)	5.3	1.9	198	70 (35.4)	0.122	0.078	26.6	16.4	39.9
Nicholaichuk (2001)	4.5	2.0	168	38 (22.6)	0.391	0.109	8.5	4.0	17.1
*Swinburne Romine et al. (2008)	2.0	1.8	439	40 (9.1)	0.292	0.080	8.2	5.9	11.3
*Ternowski (2004)	2.2	2.0	202	12 (5.9)	0.273	0.125	4.9	2.5	9.2
*Wilson et al. (2007a, b)	5.8	2.2	44	4 (9.1)	0.115	0.244	6.0	0.7	37.7

* Refers to a subsample of the full study.

Table D-15. *Completed Treatment*

	Static-99 *M*	*(SD)*	*N*	*N* recid (%)	B_1	B_1SE	B_0	B_0 95% C.I.	
No									
*Craissati et al. (2008)	2.2	1.6	66	7 (10.6)	0.400	0.244	8.6	3.6	19.2
*Haag (2005)	4.0	2.0	95	19 (20.0)	0.304	0.138	11.0	4.9	22.9
*Swinburne Romine et al. (2008)	1.9	1.8	417	42 (10.1)	0.336	0.080	8.9	6.4	12.2
*Ternowski (2004)	2.4	2.1	63	7 (11.1)	0.213	0.171	9.7	4.3	20.5
Yes									
Allan et al. (2007)	2.3	2.0	299	35 (11.7)	0.411	0.086	8.3	5.5	12.5
*Craissati et al. (2008)	2.4	2.2	134	8 (6.0)	0.402	0.141	3.6	1.4	9.0
*Haag (2005)	3.9	2.1	103	20 (19.4)	0.288	0.129	11.2	5.2	22.3
*Swinburne Romine et al. (2008)	1.7	1.6	145	6 (4.1)	-0.364	0.344	3.3	1.2	8.8
*Ternowski (2004)	2.1	1.9	184	9 (4.9)	0.252	0.153	4.3	2.1	8.6

* Refers to a subsample of the full study.

Table D-16. *Setting*

	Static-99 *M*	(*SD*)	*N*	*N* recid (%)	B_1	B_1SE	B_0	B_0 95% C.I.	
Primarily Corrections									
Allan et al. (2007)	2.3	2.0	299	35 (11.7)	0.411	0.086	8.3	5.5	12.5
Bartosh et al. (2003)	3.2	2.2	90	12 (13.3)	0.138	0.135	11.2	5.5	21.2
Bigras (2007)	2.5	1.9	207	19 (9.2)	0.362	0.117	6.4	3.6	11.2
Boer (2003)	3.3	2.3	299	11 (3.7)	0.515	0.146	1.0	0.3	3.4
Bonta & Yessine (2005)	5.2	2.0	101	19 (18.8)	0.265	0.139	8.4	2.8	22.5
Brouillette-Alarie & Proulx (2008)	3.8	2.2	199	29 (14.6)	0.360	0.095	6.7	3.5	12.5
Cortoni & Nunes (2007)	3.2	1.9	17	0 (0.0)	-	-	1.5	<0.1	43.6
Craissati et al. (2008)	2.3	2.0	200	15 (7.5)	0.358	0.114	5.4	2.8	9.9
Eher et al. (2008)	2.3	1.7	151	3 (2.0)	0.695	0.255	0.6	0.1	4.6
Endrass et al. (in press)	3.5	1.7	95	8 (8.4)	0.470	0.205	3.4	0.9	11.3
Epperson (2003)	2.8	2.2	151	16 (10.6)	0.384	0.116	6.2	3.1	12.2
Haag (2005)	3.9	2.0	198	39 (19.7)	0.296	0.094	11.1	6.4	18.5
Hanson et al. (2007)	3.1	2.1	31	0 (0.0)	-	-	0.8	<0.1	29.4
Harkins & Beech (2007)	2.8	2.2	198	19 (9.6)	0.377	0.108	5.6	2.9	10.6
Johansen (2007)	3.0	2.0	272	16 (5.9)	0.234	0.120	4.3	2.2	8.1
*Knight & Thornton (2007)	3.8	2.2	235	37 (15.7)	0.255	0.085	9.7	5.8	15.8
Långström (2004)	2.4	2.0	1278	69 (5.4)	0.328	0.054	3.9	2.9	5.2
Langton (2003)	3.3	2.1	226	23 (10.2)	0.250	0.106	6.8	3.8	12.1
Swinburne Romine et al. (2008)	1.9	1.8	570	48 (8.4)	0.279	0.073	7.8	5.8	10.4
Ternowski (2004)	2.1	1.9	247	16 (6.5)	0.246	0.114	5.7	3.3	9.5
Wilson et al. (2007a & b)	5.5	2.1	100	12 (12.0)	0.027	0.148	11.0	3.6	29.1
Primarily Mental Health									
de Vogel et al. (2004)	6.1	1.7	100	26 (26.0)	0.459	0.150	4.4	1.1	16.6
*Knight & Thornton (2007)	5.3	1.9	198	70 (35.4)	0.122	0.078	26.6	16.4	39.9
Milton (2003)	4.8	2.0	93	15 (16.1)	0.428	0.161	4.4	1.2	14.8
Nicholaichuk (2001)	4.5	2.0	168	38 (22.6)	0.391	0.109	8.5	4.0	17.1
Mixed									
Bengtson (2008)	3.8	2.1	310	61 (19.7)	0.210	0.070	13.8	9.5	19.7
Hill et al. (2008)	5.0	1.8	73	8 (11.0)	0.172	0.212	6.6	1.4	26.1

* Refers to a subsample of the full study.

Table D-17. *Sample Type*

	Static-99 *M (SD)*		*N*	*N* recid (%)	B_1	B_1SE	B_0	B_0 95% C.I.	
Preselected High Risk									
Bengtson (2008)	3.8	2.1	310	61 (19.7)	0.210	0.070	13.8	9.5	19.7
Bonta & Yessine (2005)	5.2	2.0	101	19 (18.8)	0.265	0.139	8.4	2.8	22.5
de Vogel et al. (2004)	6.1	1.7	100	26 (26.0)	0.459	0.150	4.4	1.1	16.6
Haag (2005)	3.9	2.0	198	39 (19.7)	0.296	0.094	11.1	6.4	18.5
Hill et al. (2008)	5.0	1.8	73	8 (11.0)	0.172	0.212	6.6	1.4	26.1
Knight & Thornton (2007)	4.5	2.2	433	107 (24.7)	0.241	0.054	14.4	10.2	19.9
Milton (2003)	4.8	2.0	93	15 (16.1)	0.428	0.160	4.4	1.2	14.8
Nicholaichuk (2001)	4.5	2.0	168	38 (22.6)	0.391	0.109	8.5	4.0	17.1
Wilson et al. (2007a & b)	5.5	2.1	100	12 (12.0)	0.027	0.148	11.0	3.6	29.1
Routine CSC									
Bigras (2007)	2.5	1.9	207	19 (9.2)	0.362	0.117	6.4	3.6	11.2
Boer (2003)	3.3	2.3	299	11 (3.7)	0.515	0.146	1.0	0.3	3.4
*Brouillette-Alarie & Proulx (2008)	4.4	2.2	94	15 (16.0)	0.385	0.138	5.7	1.9	15.5
Cortoni & Nunes (2007)	3.2	1.9	17	0 (0.0)	-	-	1.5	<0.1	43.6
Langton (2003)	3.3	2.1	226	23 (10.2)	0.250	0.106	6.8	3.8	12.1
Routine European Prison									
Eher et al. (2008)	2.3	1.7	151	3 (2.0)	0.695	0.255	0.6	0.1	4.6
Endrass et al. (in press)	3.5	1.7	95	8 (8.4)	0.470	0.205	3.4	0.9	11.3
Långström (2004)	2.4	2.0	1278	69 (5.4)	0.328	0.054	3.9	2.9	5.2
Community Samples									
*Brouillette-Alarie & Proulx (2008)	3.8	1.9	35	5 (14.3)	0.771	0.341	2.1	0.2	20.1
Craissati et al. (2008)	2.3	2.0	200	15 (7.5)	0.358	0.114	5.4	2.8	9.9
*Hanson et al. (2007)	2.3	1.8	18	0 (0.0)	-	-	1.4	<0.1	42.2
Swinburne Romine et al. (2008)	1.9	1.8	570	48 (8.4)	0.279	0.073	7.8	5.8	10.4

* Refers to a subsample of the full study.

Table D-18. *Sample Type Revised 1*

	Static-99 *M (SD)*		*N*	*N* recid (%)	B_1	B_1SE	B_0	B_0 95% C.I.	
Routine Corrections									
Bartosh et al. (2003)	3.2	2.2	90	12 (13.3)	0.138	0.135	11.2	5.5	21.2
Bigras (2007)	2.5	1.9	207	19 (9.2)	0.362	0.117	6.4	3.6	11.2
Boer (2003)	3.3	2.3	299	11 (3.7)	0.515	0.146	1.0	0.3	3.4
Brouillette-Alarie & Proulx (2008)	3.8	2.2	199	29 (14.6)	0.360	0.095	6.7	3.5	12.5
Cortoni & Nunes (2007)	3.2	1.9	17	0 (0.0)	-	-	1.5	<0.1	43.6
Craissati et al. (2008)	2.3	2.0	200	15 (7.5)	0.358	0.114	5.4	2.8	9.9
Eher et al. (2008)	2.3	1.7	151	3 (2.0)	0.695	0.255	0.6	0.1	4.6
Endrass et al. (in press)	3.5	1.7	95	8 (8.4)	0.470	0.205	3.4	0.9	11.3
Epperson (2003)	2.8	2.2	151	16 (10.6)	0.384	0.116	6.2	3.1	12.2
Hanson et al. (2007)	3.1	2.1	31	0 (0.0)	-	-	0.8	<0.1	29.4
Långström (2004)	2.4	2.0	1278	69 (5.4)	0.328	0.054	3.9	2.9	5.2
Langton (2003)	3.3	2.1	226	23 (10.2)	0.250	0.106	6.8	3.8	12.1
Ternowski (2004)	2.1	1.9	247	16 (6.5)	0.246	0.114	5.7	3.3	9.5
Preselected High Risk (refined definition)									
de Vogel et al. (2004)	6.1	1.7	100	26 (26.0)	0.459	0.150	4.4	1.1	16.6
Haag (2005)	3.9	2.0	198	39 (19.7)	0.296	0.094	11.1	6.4	18.5
*Knight & Thornton (2007)	5.3	1.9	198	70 (35.4)	0.122	0.078	26.6	16.4	39.9
Wilson et al. (2007a & b)	5.5	2.1	100	12 (12.0)	0.027	0.148	11.0	3.6	29.1
Remaining									
Allan et al. (2007)	2.3	2.0	299	35 (11.7)	0.411	0.086	8.3	5.5	12.5
Bengtson (2008)	3.8	2.1	310	61 (19.7)	0.210	0.070	13.8	9.5	19.7
Bonta & Yessine (2005)	5.2	2.0	101	19 (18.8)	0.265	0.139	8.4	2.8	22.5
Harkins & Beech (2007)	2.8	2.2	198	19 (9.6)	0.377	0.108	5.6	2.9	10.6
Hill et al. (2008)	5.0	1.8	73	8 (11.0)	0.172	0.212	6.6	1.4	26.1
Johansen (2007)	3.0	2.0	272	16 (5.9)	0.234	0.120	4.3	2.2	8.1
*Knight & Thornton (2007)	3.8	2.2	235	37 (15.7)	0.255	0.085	9.6	5.8	15.8
Milton (2003)	4.8	2.0	93	15 (16.1)	0.428	0.160	4.4	1.2	14.8
Nicholaichuk (2001)	4.5	2.0	168	38 (22.6)	0.391	0.109	8.5	4.0	17.1
Swinburne Romine et al. (2008)	1.9	1.8	570	48 (8.4)	0.279	0.073	7.8	5.8	10.4

* Refers to a subsample of the full study.

Table D-19. *Sample Type Revised 2*

	Static-99 *M* (*SD*)		*N*	*N* recid (%)	B_1	B_1SE	B_0	B_0 95% C.I.	
Routine Corrections									
Bartosh et al. (2003)	3.2	2.2	90	12 (13.3)	0.138	0.135	11.2	5.5	21.2
Bigras (2007)	2.5	1.9	207	19 (9.2)	0.362	0.117	6.4	3.6	11.2
Boer (2003)	3.3	2.3	299	11 (3.7)	0.515	0.146	1.0	0.3	3.4
Brouillette-Alarie & Proulx (2008)	3.8	2.2	199	29 (14.6)	0.360	0.095	6.7	3.5	12.5
Cortoni & Nunes (2007)	3.2	1.9	17	0 (0.0)	-	-	1.5	<0.1	43.6
Craissati et al. (2008)	2.3	2.0	200	15 (7.5)	0.358	0.114	5.4	2.9	9.9
Eher et al. (2008)	2.3	1.7	151	3 (2.0)	0.695	0.255	0.6	0.1	4.6
Endrass et al. (in press)	3.5	1.7	95	8 (8.4)	0.470	0.205	3.4	0.9	11.3
Epperson (2003)	2.8	2.2	151	16 (10.6)	0.384	0.116	6.2	3.1	12.2
Hanson et al. (2007)	3.1	2.1	31	0 (0.0)	-	-	0.8	<0.1	29.4
Långström (2004)	2.4	2.0	1278	69 (5.4)	0.328	0.054	3.9	2.9	5.2
Langton (2003)	3.3	2.1	226	23 (10.2)	0.250	0.106	6.8	3.8	12.0
Preselected High Risk (refined definition)									
Haag (2005)	3.9	2.0	198	39 (19.7)	0.296	0.094	11.1	6.4	18.5
*Knight & Thornton (2007)	5.3	1.9	198	70 (35.4)	0.122	0.078	26.6	16.4	39.9
Wilson et al. (2007a & b)	5.5	2.1	100	12 (12.0)	0.027	0.148	11.0	3.6	29.1
Remaining									
Allan et al. (2007)	2.3	2.0	299	35 (11.7)	0.411	0.086	8.3	5.5	12.5
Bengtson (2008)	3.8	2.1	310	61 (19.7)	0.210	0.070	13.8	9.5	19.7
Bonta & Yessine (2005)	5.2	2.0	101	19 (18.8)	0.265	0.139	8.4	2.8	22.5
de Vogel et al. (2004)	6.1	1.7	100	26 (26.0)	0.459	0.150	4.4	1.1	16.6
Harkins & Beech (2007)	2.8	2.2	198	19 (9.6)	0.377	0.108	5.6	2.9	10.6
Hill et al. (2008)	5.0	1.8	73	8 (11.0)	0.172	0.212	6.6	1.4	26.1
Johansen (2007)	3.0	2.0	272	16 (5.9)	0.234	0.120	4.3	2.2	8.1
*Knight & Thornton (2007)	3.8	2.2	235	37 (15.7)	0.255	0.085	9.6	5.8	15.8
Milton (2003)	4.8	2.0	93	15 (16.1)	0.428	0.160	4.4	1.2	14.8
Nicholaichuk (2001)	4.5	2.0	168	38 (22.6)	0.391	0.109	8.5	4.0	17.1
Swinburne Romine et al. (2008)	1.9	1.8	570	48 (8.4)	0.279	0.073	7.8	5.8	10.4
Ternowski (2004)	2.1	1.9	247	16 (6.5)	0.246	0.114	5.7	3.3	9.5

* Refers to a subsample of the full study.

Table D-20. *Routine Corrections*

	Static-99 *M*	*(SD)*	*N*	*N* recid (%)	B_1	B_1SE	B_0	B_0 95% C.I.	
Routine Corrections									
Bartosh et al. (2003)	3.2	2.2	90	12 (13.3)	0.138	0.135	11.2	5.5	21.2
Bigras (2007)	2.5	1.9	207	19 (9.2)	0.362	0.117	6.4	3.6	11.2
Boer (2003)	3.3	2.3	299	11 (3.7)	0.515	0.146	1.0	0.3	3.4
Craissati et al. (2008)	2.3	2.0	200	15 (7.5)	0.358	0.114	5.4	2.8	9.9
Eher et al. (2008)	2.3	1.7	151	3 (2.0)	0.695	0.255	0.6	0.1	4.6
Endrass et al. (in press)	3.5	1.7	95	8 (8.4)	0.470	0.205	3.4	0.9	11.3
Epperson (2003)	2.8	2.2	151	16 (10.6)	0.384	0.116	6.2	3.1	12.2
Hanson et al. (2007)	3.1	2.1	31	0 (0.0)	-	-	0.8	<0.1	29.4
Långström (2004)	2.4	2.0	1278	69 (5.4)	0.328	0.054	3.9	2.9	5.2
Non-Routine									
Allan et al. (2007)	2.3	2.0	299	35 (11.7)	0.411	0.086	8.3	5.5	12.5
Brouillette-Alarie & Proulx (2008)	3.8	2.2	199	29 (14.6)	0.360	0.095	6.7	3.5	12.5
Bengtson (2008)	3.8	2.1	310	61 (19.7)	0.210	0.070	13.8	9.5	19.7
Bonta & Yessine (2005)	5.2	2.0	101	19 (18.8)	0.265	0.139	8.4	2.8	22.5
Cortoni & Nunes (2007)	3.2	1.9	17	0 (0.0)	-	-	1.5	<0.1	43.6
de Vogel et al. (2004)	6.1	1.7	100	26 (26.0)	0.459	0.150	4.4	1.1	16.6
Haag (2005)	3.9	2.0	198	39 (19.7)	0.296	0.094	11.1	6.4	18.5
Harkins & Beech (2007)	2.8	2.2	198	19 (9.6)	0.377	0.108	5.6	2.9	10.6
Hill et al. (2008)	5.0	1.8	73	8 (11.0)	0.172	0.212	6.6	1.4	26.1
Johansen (2007)	3.0	2.0	272	16 (5.9)	0.234	0.120	4.3	2.2	8.1
Knight & Thornton (2007)	4.5	2.2	433	107 (24.7)	0.241	0.054	14.4	10.2	19.9
Langton (2003)	3.3	2.1	226	23 (10.2)	0.250	0.106	6.8	3.8	12.1
Milton (2003)	4.8	2.0	93	15 (16.1)	0.428	0.160	4.4	1.2	14.8
Nicholaichuk (2001)	4.5	2.0	168	38 (22.6)	0.391	0.109	8.5	4.0	17.1
Swinburne Romine et al. (2008)	1.9	1.8	570	48 (8.4)	0.279	0.073	7.8	5.8	10.4
Ternowski (2004)	2.1	1.9	247	16 (6.5)	0.246	0.114	5.7	3.3	9.5
Wilson et al. (2007a & b)	5.5	2.1	100	12 (12.0)	0.027	0.148	11.0	3.6	29.1

Table D-21. *Routine and treatment types*

	Static-99 *M*	*(SD)*	*N*	*N* recid (%)	B_1	B_1SE	B_0	B_0 *95% C.I.*	
Routine Corrections									
Bartosh et al. (2003)	3.2	2.2	90	12 (13.3)	0.138	0.134	11.2	5.5	21.2
Bigras (2007)	2.5	1.9	207	19 (9.2)	0.362	0.117	6.4	3.6	11.2
Boer (2003)	3.3	2.3	299	11 (3.7)	0.515	0.146	1.0	0.3	3.4
Craissati et al. (2008)	2.3	2.0	200	15 (7.5)	0.358	0.114	5.4	2.8	9.9
Eher et al. (2008)	2.3	1.7	151	3 (2.0)	0.695	0.255	0.6	0.1	4.6
Endrass et al. (in press)	3.5	1.7	95	8 (8.4)	0.47	0.205	3.4	0.9	11.3
Epperson (2003)	2.8	2.2	151	16 (10.6)	0.384	0.116	6.2	3.1	12.2
Hanson et al. (2007)	3.1	2.1	31	0 (0.0)	-	-	0.8	<0.1	29.4
Långström (2004)	2.4	2.0	1278	69 (5.4)	0.328	0.054	3.9	2.9	5.2
Moderate Treatment									
Allan et al. (2007)	2.3	2.0	299	35 (11.7)	0.411	0.086	8.3	5.5	12.5
Harkins & Beech (2007)	2.8	2.2	198	19 (9.6)	0.377	0.108	5.6	2.9	10.6
Langton (2003)	3.3	2.0	226	23 (10.2)	0.250	0.106	6.8	3.8	12.1
Swinburne Romine et al. (2008)	1.9	1.8	570	48 (8.4)	0.279	0.073	7.8	5.8	10.4
Ternowski (2004)	2.1	1.9	247	16 (6.5)	0.246	0.114	5.7	3.3	9.5
Moderate-High Intensity Treatment									
Brouillette-Alarie & Proulx (2008)	3.8	2.2	199	29 (14.6)	0.360	0.095	6.7	3.5	12.5
de Vogel et al. (2004)	6.1	1.7	100	26 (26.0)	0.459	0.150	4.4	1.1	16.6
Johansen (2007)	3.0	2.0	272	16 (5.9)	0.234	0.120	4.3	2.2	8.1
Knight & Thornton (2007)	4.5	2.2	433	107 (24.7)	0.241	0.054	14.4	10.2	19.9
Nicholaichuk (2001)	4.5	2.0	168	38 (22.6)	0.391	0.109	8.5	4.0	17.1

Table D-22. *Preselected*

	Static-99 *M (SD)*		*N*	*N* recid (%)	B_1	B_1SE	B_0	B_0 95% C.I.	
Routine Corrections									
Bartosh et al. (2003)	3.2	2.2	90	12 (13.3)	0.138	0.134	11.2	5.5	21.2
Bigras (2007)	2.5	1.9	207	19 (9.2)	0.362	0.117	6.4	3.6	11.2
Boer (2003)	3.3	2.3	299	11 (3.7)	0.515	0.146	1.0	0.3	3.4
Craissati et al. (2008)	2.3	2.0	200	15 (7.5)	0.358	0.114	5.4	2.8	9.9
Eher et al. (2008)	2.3	1.7	151	3 (2.0)	0.695	0.255	0.6	0.1	4.6
Endrass et al. (in press)	3.5	1.7	95	8 (8.4)	0.47	0.205	3.4	0.9	11.3
Epperson (2003)	2.8	2.2	151	16 (10.6)	0.384	0.116	6.2	3.1	12.2
Hanson et al. (2007)	3.1	2.1	31	0 (0.0)	-	-	0.8	<0.1	29.4
Långström (2004)	2.4	2.0	1278	69 (5.4)	0.328	0.054	3.9	2.9	5.2
Preselect Treatment									
Allan et al. (2007)	2.3	2.0	299	35 (11.7)	0.411	0.086	8.3	5.5	12.5
Brouillette-Alarie & Proulx (2008)	3.8	2.2	199	29 (14.6)	0.360	0.095	6.7	3.5	12.5
Harkins & Beech (2007)	2.8	2.2	198	19 (9.6)	0.377	0.108	5.6	2.9	10.6
Johansen (2007)	3.0	2.0	272	16 (5.9)	0.234	0.120	4.3	2.2	8.1
Langton (2003)	3.3	2.1	226	23 (10.2)	0.250	0.106	6.8	3.8	12.1
Swinburne Romine et al. (2008)	1.9	1.8	570	48 (8.4)	0.279	0.073	7.8	5.8	10.4
Ternowski (2004)	2.1	1.9	247	16 (6.5)	0.246	0.114	5.7	3.3	9.5
Preselect Risk									
Bonta & Yessine (2005)	5.2	2.0	101	19 (18.8)	0.265	0.139	8.4	2.8	22.5
de Vogel et al. (2004)	6.1	1.7	100	26 (26.0)	0.459	0.150	4.4	1.1	16.6
Haag (2005)	3.9	2.0	198	39 (19.7)	0.296	0.094	11.1	6.4	18.5
Knight & Thornton (2007)	4.5	2.2	433	107 (24.7)	0.241	0.054	14.4	10.2	19.9
Nicholaichuk (2001)	4.5	2.0	168	38 (22.6)	0.391	0.109	8.5	4.0	17.1
Wilson et al. (2007a & b)	5.5	2.1	100	12 (12.0)	0.027	0.148	11.0	3.6	29.1
Preselect Psychiatric									
Bengtson (2008)	3.8	2.1	310	61 (19.7)	0.210	0.070	13.8	9.5	19.7
Milton (2003)	4.8	2.0	93	15 (16.1)	0.428	0.160	4.4	1.2	14.8

Table D-23. *Preselected Version 2*

	Static-99 *M*	*(SD)*	*N*	*N* recid (%)	B_1	B_1SE	B_0	B_0 95% C.I.	
Routine Corrections									
Bartosh et al. (2003)	3.2	2.2	90	12 (13.3)	0.138	0.134	11.2	5.5	21.2
Bigras (2007)	2.5	1.9	207	19 (9.2)	0.362	0.117	6.4	3.6	11.2
Boer (2003)	3.3	2.3	299	11 (3.7)	0.515	0.146	1.0	0.3	3.4
Craissati et al. (2008)	2.3	2.0	200	15 (7.5)	0.358	0.114	5.4	2.8	9.9
Eher et al. (2008)	2.3	1.7	151	3 (2.0)	0.695	0.255	0.6	0.1	4.6
Endrass et al. (in press)	3.5	1.7	95	8 (8.4)	0.47	0.205	3.4	0.9	11.3
Epperson (2003)	2.8	2.2	151	16 (10.6)	0.384	0.116	6.2	3.1	12.2
Hanson et al. (2007)	3.1	2.1	31	0 (0.0)	-	-	0.8	<0.1	29.4
Långström (2004)	2.4	2.0	1278	69 (5.4)	0.328	0.054	3.9	2.9	5.2
Preselect Treatment									
Allan et al. (2007)	2.3	2.0	299	35 (11.7)	0.411	0.086	8.3	5.5	12.5
Brouillette-Alarie & Proulx (2008)	3.8	2.2	199	29 (14.6)	0.360	0.095	6.7	3.5	12.5
Harkins & Beech (2007)	2.8	2.2	198	19 (9.6)	0.377	0.108	5.6	2.9	10.6
Johansen (2007)	3.0	2.0	272	16 (5.9)	0.234	0.120	4.3	2.2	8.1
Langton (2003)	3.3	2.1	226	23 (10.2)	0.250	0.106	6.8	3.8	12.1
Swinburne Romine et al. (2008)	1.9	1.8	570	48 (8.4)	0.279	0.073	7.8	5.8	10.4
Ternowski (2004)	2.1	1.9	247	16 (6.5)	0.246	0.114	5.7	3.3	9.5
Preselect Risk/Need									
Bengtson (2008)	3.8	2.1	310	61 (19.7)	0.210	0.070	13.8	9.5	19.7
Bonta & Yessine (2005)	5.2	2.0	101	19 (18.8)	0.265	0.139	8.4	2.8	22.5
de Vogel et al. (2004)	6.1	1.7	100	26 (26.0)	0.459	0.150	4.4	1.1	16.6
Haag (2005)	3.9	2.0	198	39 (19.7)	0.296	0.094	11.1	6.4	18.5
Knight & Thornton (2007)	4.5	2.2	433	107 (24.7)	0.241	0.054	14.4	10.2	19.9
Milton (2003)	4.8	2.0	93	15 (16.1)	0.428	0.160	4.4	1.2	14.8
Nicholaichuk (2001)	4.5	2.0	168	38 (22.6)	0.391	0.109	8.5	4.0	17.1
Wilson et al. (2007a & b)	5.5	2.1	100	12 (12.0)	0.027	0.148	11.0	3.6	29.1

Table D-24. *Preselected Version 3*

	Static-99 *M*	(*SD*)	*N*	*N* recid (%)	B_1	B_1SE	B_0	B_0 95% C.I.	
Routine Corrections (Non-U.S.)									
Bigras (2007)	2.5	1.9	207	19 (9.2)	0.362	0.117	6.4	3.6	11.2
Boer (2003)	3.3	2.3	299	11 (3.7)	0.515	0.146	1.0	0.3	3.4
Craissati et al. (2008)	2.3	2.0	200	15 (7.5)	0.358	0.114	5.4	2.8	9.9
Eher et al. (2008)	2.3	1.7	151	3 (2.0)	0.695	0.255	0.6	0.1	4.6
Endrass et al. (in press)	3.5	1.7	95	8 (8.4)	0.470	0.205	3.4	0.9	11.3
Hanson et al. (2007)	3.1	2.1	31	0 (0.0)	-	-	0.8	<0.1	29.4
Långström (2004)	2.4	2.0	1278	69 (5.4)	0.328	0.054	3.9	2.9	5.2
Routine Corrections (U.S.)									
Bartosh et al. (2003)	3.2	2.2	90	12 (13.3)	0.138	0.135	11.2	5.5	21.2
Epperson (2003)	2.8	2.2	151	16 (10.6)	0.384	0.116	6.2	3.1	12.2
Preselect Treatment									
Allan et al. (2007)	2.3	2.0	299	35 (11.7)	0.411	0.086	8.3	5.5	12.5
Brouillette-Alarie & Proulx (2008)	3.8	2.2	199	29 (14.6)	0.360	0.095	6.7	3.5	12.5
Harkins & Beech (2007)	2.8	2.2	198	19 (9.6)	0.377	0.108	5.6	2.9	10.6
Johansen (2007)	3.0	2.0	272	16 (5.9)	0.234	0.120	4.3	2.2	8.1
Langton (2003)	3.3	2.2	226	23 (10.2)	0.250	0.106	6.8	3.8	12.1
Swinburne Romine et al. (2008)	1.9	1.8	570	48 (8.4)	0.279	0.073	7.8	5.8	10.4
Ternowski (2004)	2.1	1.9	247	16 (6.5)	0.246	0.114	5.7	3.3	9.5
Preselect Risk/Need									
Bengtson (2008)	3.8	2.1	310	61 (19.7)	0.210	0.070	13.8	9.5	19.7
Bonta & Yessine (2005)	5.2	2.0	101	19 (18.8)	0.265	0.139	8.4	2.8	22.5
de Vogel et al. (2004)	6.1	1.7	100	26 (26.0)	0.459	0.150	4.4	1.1	16.6
Haag (2005)	3.9	2.0	198	39 (19.7)	0.296	0.094	11.1	6.4	18.5
Knight & Thornton (2007)	4.5	2.2	433	107 (24.2)	0.241	0.054	14.4	10.2	19.9
Milton (2003)	4.8	2.0	93	15 (16.1)	0.428	0.160	4.4	1.2	14.8
Nicholaichuk (2001)	4.5	2.0	168	38 (22.6)	0.391	0.109	8.5	4.0	17.1
Wilson et al. (2007a & b)	5.5	2.1	100	12 (12.0)	0.027	0.148	11.0	3.6	29.1

Appendix E

Static-99R logistic regression analyses at 5 years

	N	*N* recid (%)	B_1	B_1SE	$B_{0(2)}$	$B_{0(2)}SE$
Allan et al. (2007)	298	35 (11.7)	0.439	0.087	-2.332	0.228
Bartosh et al. (2007)	90	12 (13.3)	0.137	0.114	-2.054	0.367
Bengtson (2008)	310	61 (19.7)	0.208	0.065	-1.851	0.216
Bigras (2007)	207	19 (9.2)	0.354	0.111	-2.577	0.298
Boer (2003)	299	11 (3.7)	0.467	0.140	-4.364	0.596
Bonta & Yessine (2005)	101	19 (18.8)	0.347	0.144	-2.698	0.628
Brouillette-Alarie & Proulx (2008)	199	29 (14.6)	0.332	0.092	-2.601	0.352
Cortoni & Nunes (2007)	17	0 (0.0)	-	-	-4.205	2.015
Craissati et al. (2008)	200	15 (7.5)	0.341	0.112	-2.820	0.332
Eher et al. (2008)	151	3 (2.0)	1.013	0.401	-5.801	1.522
Epperson (2003)	150	16 (10.7)	0.347	0.107	-2.651	0.371
Haag (2005)	198	39 (19.7)	0.299	0.090	-2.120	0.310
Hanson et al. (2007)	31	0 (0.0)	-	-	-4.812	2.008
Harkins & Beech (2007)	197	19 (9.6)	0.350	0.097	-2.648	0.323
Hill et al. (2008)	73	8 (11.0)	0.377	0.232	-3.340	0.948
Johansen (2007)	272	16 (5.9)	0.208	0.111	-3.056	0.332
Knight & Thornton (2007)	433	107 (24.7)	0.240	0.052	-1.806	0.202
Långström (2004)	1278	69 (5.4)	0.308	0.050	-3.118	0.151
Nicholaichuk (2001)	168	38 (22.6)	0.362	0.099	-2.476	0.430
Swinburne Romine et al. (2008)	569	48 (8.4)	0.249	0.066	-2.442	0.160
Ternowski (2004)	247	16 (6.5)	0.296	0.100	-2.786	0.291
Wilson et al. (2007a & b)	103	12 (11.7)	0.035	0.136	-2.139	0.545

Appendix F

Static-99R logistic regression analyses at 10 years

	N	*N* recid	B_1	B_1SE	$B_{0(2)}$	$B_{0(2)}SE$
Allan et al. (2007)	25	5	0.900	0.419	-1.8632	0.673
Bengtson (2008)	291	83	0.230	0.061	-1.398	0.195
Boer (2003)	295	23	0.378	0.094	-3.211	0.358
Brouillette-Alarie & Proulx (2008)	110	23	0.261	0.106	-1.930	0.371
Craissati et al. (2008)	66	6	0.200	0.182	-2.262	0.434
Epperson (2003)	36	8	0.794	0.290	-3.650	1.237
Harkins & Beech (2007)	127	21	0.362	0.102	-2.005	0.311
Hill et al. (2008)	54	10	0.230	0.207	-2.230	0.802
Johansen (2007)	62	8	0.067	0.165	-2.024	0.485
Knight & Thornton (2007)	353	106	0.198	0.051	-1.371	0.189
Långström (2004)	353	26	0.407	0.086	-2.986	0.280
Nicholaichuk (2001)	59	15	0.234	0.143	-1.908	0.630
Swinburne Romine et al. (2008)	542	61	0.213	0.061	-2.088	0.141
Wilson et al. (2007a & b)	16	1	33.44	4694.13	-217.62	29918.49

Appendix G

Recidivism Tables for Static-99R (Recidivism is Reported as a Percentage)

	5 Years			10 Years			
Score	Routine	Treatment Need	Risk	Routine - observed	Routine - adjusted	Treatment Need	Risk
-3	1.2	1.8		1.4	1.8	3.2	
-2	1.6	2.4		1.8	2.4	4.2	
-1	2.2	3.2	5.5	2.3	3.3	5.4	9.8
0	2.9	4.2	7.2	3.0	4.4	7.0	12.5
1	3.8	5.5	9.4	3.9	5.7	9.0	15.7
2	5.0	7.2	12.2	5.1	7.6	11.5	19.7
3	6.6	9.4	15.7	6.6	10.0	14.5	24.3
4	8.6	12.2	19.9	8.4	13.0	18.2	29.6
5	11.2	15.6	24.9	10.8	16.9	22.6	35.5
6	14.4	19.9	30.7	13.7	21.7	27.6	41.9
7	18.4	24.9	37.2	17.2	27.8	33.3	48.6
8	23.2	30.7	44.2	21.4	35.0	39.6	55.3
9	28.7	37.2	51.4	26.3	43.3	46.2	61.9
10			58.6				68.0

Appendix H

Equations for calculating Static-99R recidivism percentages (in SPSS syntax)

*Note that in regression, the predicted value on a dependent variable (recidivism) for a given score on the independent variable (Static-99R) is obtained by adding the intercept (the predicted value for a score of 0) to the product of the slope and the independent variable. In other words, $y = B_0 + x*B_1$

*B_0 was calculated with Static-99R centered on a score of 2. Therefore, in the syntax below, the independent variable (Static99R2) refers to the Static-99R score minus 2.

```
COMPUTE Routine5yr = -2.939181 + Static99R2*0.289910.
COMPUTE rout5perc=(EXP(Routine5yr) / (1+EXP(Routine5yr)))*100.
EXECUTE.

COMPUTE Treatment5yr = -2.553833 + Static99R2*0.289910.
COMPUTE treat5perc=(EXP(Treatment5yr) / (1+EXP(Treatment5yr)))*100.
EXECUTE.

COMPUTE RiskPsych5yr = -1.971812 + Static99R2*0.289910.
COMPUTE risk5perc=(EXP(RiskPsych5yr) / (1+EXP(RiskPsych5yr)))*100.
EXECUTE.

COMPUTE Routine10yr = -2.923503 + Static99R2*0. 270301.
COMPUTE rout10perc=(EXP(Routine10yr) / (1+EXP(Routine10yr)))*100.
EXECUTE.

COMPUTE Treatment10yr = -2.044199 + Static99R2*0. 270301.
COMPUTE treat10perc=(EXP(Treatment10yr) / (1+EXP(Treatment10yr)))*100.
EXECUTE .

COMPUTE RiskPsych10yr = -1.407027 + Static99R2*0. 270301.
COMPUTE risk10perc=(EXP(RiskPsych10yr) / (1+EXP(RiskPsych10yr)))*100.
EXECUTE.
```