Reporting Quality of the Sex Offender Risk Assessment Literature: A Systematic Review

Cornel Gmür[1,2], Jay P. Singh[2,3,4]
[1] Psychiatric-Psychological Service, Department of Justice, Canton of Zurich, Switzerland
[2] Department of Psychology, University of Konstanz, Konstanz, Germany
[3] Global Institute of Forensic Research, Reston, VA, United States of America
[4] Faculty of Health Sciences, Molde University College, Molde, Norway

Abstract

The aim of the present study was to investigate the reporting quality of risk assessment predictive validity studies published in leading sex- and gender-related journals according to the Risk Assessment Guidelines for the Evaluation of Efficacy (RAGEE) Statement checklist. A systematic search for sex- and gender-related journals identified two that had published at least five articles investigating the predictive validity of a risk assessment method between 2008 and 2013: Sexual Abuse: A Journal of Research and Treatment (SAJRT; k articles = 21) and Journal of Sexual Aggression (JSA; k = 6). A RAGEE Statement checklist was coded for each article, and both sum score and item-level analyses were conducted. An average of 39.5 (SD = 3.0, range = 34-46) of the 50 RAGEE criteria were satisfied, suggesting that there is 20% room for improvement. SAJRT studies met significantly more RAGEE criteria than JSA studies. However, important sample- and study-level information needed to interpret findings was frequently missing across articles. Adopting RAGEE Statement guidelines into Instructions for Authors in journals that publish predictive validity studies on risk assessment methods for sex offenders has the potential to improve the quality and consistency of this literature.

Keywords: RAGEE Statement, reporting quality, sexual offender, risk assessment, recidivism

Introduction

In recent years, the use of structured risk assessment tools to assess sex offender recidivism risk has been increasing in criminal justice systems, mental health services, and the interface between them. In Western countries including the United States, United Kingdom, Canada, Australia, and Switzerland such instruments are used to aide in decisions regarding individual liberty and public protection (Helmus, Hanson, & Morton-Bourgon, 2011; Rossegger, Gerth, Urbaniok, Laubacher, & Endrass, 2010). Supporting the use of such instruments is a large evidence base on predictive validity that has grown exponentially in recent years (Singh, Fazel, Gueorguieva, & Buchanan, 2013).

A recent systematic review by Singh, Desmarais, and Van Dorn (2013) investigated the quality of methodological reporting in predictive validity studies of a number of structured instruments, including the following sex offender tools: Minnesota Sex Offender Screening Tool-Revised (Epperson et al., 1998), Rapid Risk Assessment for Sexual Offense Recidivism (Hanson, 1997), Risk Matrix 2000 (Thornton et al., 2003), Sex Offender Risk Appraisal Guide (Quinsey, Harris, Rice, & Cormier, 2006), Static-99 (Hanson & Thornton, 1999), Static-2002 (Hanson & Thornton, 2003), and Sexual Violence Risk-20 (Boer, Hart, Kropp, & Webster, 1997). The review concluded that "measurement practices varied considerably across articles" (p. 66), with key pieces of

methodological information frequently going unreported (e.g., whether predictive validity estimates were based on total scores or categorical risk judgments).

To address this, the first set of reporting guidance for risk assessment studies, the Risk Assessment Guidelines for the Evaluation of Efficacy (RAGEE) Statement, was recently published (Singh, Yang, Mulvey, & the RAGEE Group, in press). This 50-item reporting checklist was developed by a multidisciplinary panel of 37 international experts in the field of risk assessment from 10 countries. The panelists endorsed the RAGEE Statement checklist as being highly satisfactory and as indicating study features that should be reported routinely in risk assessment manuscripts. The innovative developments of the RAGEE Group appear promising, but in order to evaluate whether reporting practices are improving over time, it is first necessary to establish a baseline level of transparency in the current literature.

The Present Study

The aim of the present review was to systematically assess the reporting quality of the sex offender risk assessment research literature in leading sex- and gender-related journals using the RAGEE Statement. Specifically, we aimed to establish a baseline estimate of the transparency achieved in the current literature, as well as to investigate whether such transparency fluctuates over time. Finally, we wished to examine differences in transparency by journal both descriptively as well as statistically. This review represents the first use of the RAGEE Statement to examine the transparency and consistency with which methodology and findings are reported in the sex offender risk assessment literature.

Method

Systematic Search

Sex- and gender-related journals were identified using a recently-published, comprehensive bibliography (Zucker, 2013) supplemented by consultation with experts in the field of sex offender risk assessment. Seventy-eight eligible journals were identified (see Supplement 1 for full list). The titles and abstracts of primary studies published between January 1, 2008 and November 1, 2013 were screened for evidence of predictive validity analyses having been conducted for either unstructured clinical judgments or use of a risk assessment tool to evaluate recidivism risk.[1] Journals were excluded if fewer than five articles including such predictive validity analyses had been published in the target timeframe, as reporting quality was not able to be compared using small-sample comparisons of exact levels (Larntz, 1978). Using this strategy, *Sexual Abuse: A Journal of Research and Treatment* (*SAJRT; k* articles = 21) and *Journal of Sexual Aggression* (*JSA: k* = 6) met inclusion criteria and underwent subsequent data extraction and analytic procedures. The included *SAJRT* articles included the following work: Brouillette-Alarie & Proulx, 2012; Craissati, Bierer, & South, 2010; Griffin, Beech, Print, Bradshaw, & Quayle, 2008; Griffin & Vettor, 2011; Skelton & Vess, 2008; Wilcox, Beech, Markall, & Blacker, 2009. The JSA articles included the following work: Aebi, Plattner, Steinhausen, & Bessler, 2011; Barnett, Wakeling, & Howard, 2010; Beggs & Grace, 2010; Brown, Harkins, & Beech, 2012; Chu, Ng, Fong, & Teoh, 2012; Duwe & Freske, 2012; Eher, Matthes, Schilling, Haubner-MacLean, & Rettenberger, 2012; Fanniff & Letourneau, 2012; Grady, Edwards, Pettus-Davis, & Abramson, 2013; Grubin, 2011; Harkins, Beech, & Goodwill, 2010; Kingston, Yates, Firestone, Babchishin, & Bradford, 2008; McGrath, Lasher, & Cumming, 2012; Montana et al., 2012; Olver, Nicholaichuk, Gu, & Wong, 2013; Smallbone & Rallings, 2013; Storey, Watt, Jackson, & Hart, 2012; Swinburne Romine, Miner, Poulin, Dwyer, & Berg, 2012; Wakeling, Howard, & Barnett, 2011; Wilson, Abracen, Looman, Picheca, & Ferguson, 2011; Worling, Bookalam, & Litteljohn, 2012).

Data Extraction

The first author coded a RAGEE Statement checklist for each of the 27 predictive validity articles. If a checklist criterion was satisfied, it was coded as a "1", and if it was not, as a "0". If a checklist criterion was not satisfied because it was not applicable to the sample, study design, or instrument under investigation, it was coded as a "1" for the purposes of this study. When more than one risk assessment method was investigated (e.g., the predictive validity of two instruments was compared) but a RAGEE criterion was only met for one of those methods (e.g., inter-rater reliability measured for one instrument but not another), then the criterion was coded as a "0" for the purposes of this study. When an article stated that methodological details were available in a previous manuscript, relevant RAGEE criteria were extracted from that previous manuscript.

As a measure of quality control, three (11.1%) of the included articles were randomly selected and coded by the corresponding author. A high level of inter-rater agreement was established (κ = 0.81; Landis & Koch, 1977). Discrepancies were resolved upon discussion.

Statistical Analyses

The proportion of articles that met each RAGEE criteria was descriptively examined both overall as well as for *SAJRT* and *JSA*, separately. In addition, Fisher's (1922) exact test, a statistical significance test measuring the association between two classification methods with small sample sizes, was used to investigate differences in proportions between the number of *SAJRT* articles that met each RAGEE criterion compared to *JSA* articles. Next, a Student's t-test was conducted to evaluate the difference in the overall number of RAGEE criteria met in *SAJRT* articles compared to *JSA* articles. Finally, a correlation coefficient (Spearman's ρ) was calculated to test the presence and magnitude of a relationship between the number of RAGEE criteria satisfied and the date of publication. All analyses were conducted in MedCalc Version 11.3.8.0 for Windows using a standard significant threshold of α = 0.05.

Results

Descriptive Characteristics

The average article fulfilled 39.5 (SD = 3.0, range = 34-46) of the 50 RAGEE criteria. An average of 2.7 (SD = 0.5, range = 1-3) of the four Abstract section criteria were met, 2.0 (SD = 0.0, range = 2-2) of the two Introduction section criteria were met, 22.0 (SD = 2.6, range = 17-27) of the 30 Method section criteria were met, 5.4 (SD = 0.6, range = 4-6) of the six Results section criteria were met, 3.5 (SD = 0.6, range = 2-4) of the four Discussion section criteria were met, and 3.9 (SD = 0.2, range = 3-4) of the four Disclosure section criteria were met. On average, the "Study Design" subsection of the Methods had the most criteria fulfilled (M = 4.0 of 5 criteria, SD = 0.9, range = 2-5). The "Predicted Outcome" subsection of the Methods had the fewest criteria met (M = 1.3 of 3 criteria, SD = 0.6, range = 0-2). The articles that satisfied the most RAGEE criteria included the works of Barnett and colleagues (2010; *n* criteria satisfied = 46, 92.0%), Faniff and colleagues (2012; n = 44, 88.0%), and Worling and colleagues (2012; n = 44, 88.0%).

Association Findings

When differences in the proportion of articles meeting each RAGEE criterion were calculated between the two journals, Fisher's exact tests revealed that *SAJRT* articles were significantly more likely to: (1) identify the risk assessment instrument(s) whose predictive validity is measured in their

Abstracts, (2) report the acronym(s) and full name(s) of the instrument(s) under investigation with appropriate referencing in their Methods, (3) report the population for which the instrument(s) under investigation was intended to be used in their Methods, (4) identify the statistical significance level used in their Methods, and (5) discuss the generalizability of study findings in their Discussions, $^2(1, N = 27) \geq 3.06$, $p < 0.05$. *SAJRT* articles ($M_n = 40.4$, $SD = 2.7$) met significantly more RAGEE criteria than *JSA* articles ($M_n = 36.8$, $SD = 2.8$), $t(25) = 2.86$, $p < 0.01$. There was no association between date of publication and the total RAGEE score, (25) = -0.03, $p = 0.87$.

Discussion

The present study represents the first systematic review of the reporting quality achieved in the current sex offender risk assessment literature using criteria set forth by the RAGEE Group. A systematic search identified two journals that had published at least five articles investigating the predictive validity of a risk assessment method between 2008 and 2013: *Sexual Abuse: A Journal of Research and Treatment and Journal of Sexual Aggression.* A RAGEE Statement checklist was coded for each article, and both sum score and item-level analyses were conducted. An average of 40 of the 50 RAGEE criteria were satisfied, implying an approximately 20% room for improvement. And this may be an overestimate, as it was assumed that not undisclosed competing interests did not exist, whereas recent systematic review evidence suggests that this may not always be the case (Singh, Grann, & Fazel, 2013). In addition, the majority of studies did not report whether modifications were made to manual-based tool administration protocols, so it was assumed that no modifications were made. However, review evidence also exists that this is likely an overly conservative assumption (Rossegger et al., 2013). Overall, SAJRT articles were found to satisfy more RAGEE items than JSA, and no association was found between the date of publication and the number of RAGEE criteria satisfied, suggesting that the reporting quality of the sex offender risk assessment literature has remained stable over most of the past decade. It is important to note that the aim of this study (and the RAGEE initiative more generally) was not to investigate whether methodology and findings were reported correctly, but rather whether critical information in these areas was reported at all.

Implications

The results of the present review have potentially important implications for researchers and practitioners. First, the finding that fundamental pieces of information on study design are frequently absent from published manuscripts supports the adoption of the RAGEE Statement as part of the *Instructions for Authors* for journals that publish research on sex offender risk assessment. This would serve as a quality control measure and is supported by leading researchers, practitioners, legal professionals, and journal editors. Second, readers of the current research literature on sex offender risk assessment may find it difficult to objectively assess the accuracy of study findings, as potentially important sample- and study-level information is frequently missing. For example, we found that how participants were recruited, their average age at assessment, when they were assessed, their index offense composition, and characteristics of the individuals who assessed them was not commonly reported. In addition, critical descriptive information on how many participants were classified into different risk levels, and how many of those persons went on to engage in the outcome of interest was reported in a minority of articles. Practitioners are advised that caution is warranted in interpreting the strength of reliability and validity research findings in the absence of transparency (cf. Fava, 2007; Maj, 2005). Third, we have identified three articles that met the most RAGEE criteria that could be used as models for the predictive validity literature on sex offender risk assessment (Barnett et al., 2010; Faniff et al., 2012; Worling et al., 2012). Though these articles did not satisfy all 50 RAGEE conditions, they did fulfill over 85%.

Limitations and Future Directions

The present review also has several limitations. First, articles published in journals not focused on issues of gender and sexuality did not meet inclusion criteria for the present review (e.g., *Law and Human Behavior, Journal of Interpersonal Violence, Criminal Justice and Behavior*). However, recent systematic reviews have found that predictive validity studies of sex offender risk assessment have been published in these more general forensic journals (Helmus, Hanson, Thornton, Babchishin, & Harris, 2012; Singh, Grann, & Fazel, 2011). Hence, future research on the quality of reporting in risk assessment studies should explore literature published in more general journals. Second, several eligible journals with predictive validity studies in the target timeframe were excluded (e.g., *Sexual Offender Treatment and Zeitschrift für Sexualforschung*), as fewer than five such articles were published between 2008 and 2013. This was necessary to conform to established statistical standards for the use of exact tests (Larntz, 1978), but further descriptive research may assist the esteemed editorial boards of these excluded journals in identifying areas of strength and potential areas for improvement in the reporting quality of their accepted manuscripts. Third, only peer-reviewed articles published between 2008 and 2013 were considered for inclusion. Albeit the aim of this review was to establish a baseline of reporting quality for the current literature on sex offender risk assessment, it should be kept in mind that influential research on this topic has been available for a number of decades (for a review see Hall, 1990). Future research may compare the reporting quality of articles on sex offender risk assessment across time to examine trends. Given these limitations, our findings should be viewed as preliminary observations using a cross-section of the contemporary literature on sex offender risk assessment.

Conclusion

Study quality has been shown to account for differences in research findings (Rutjes et al., 2006), and it is difficult to compare study quality without transparent and consistent reporting of methodology. The RAGEE Statement is the first set of expert reporting guidance set forth for the risk assessment literature to aid in this. The findings of the present study suggest that the adoption of standardized reporting guidelines such as the RAGEE Statement in journals that publish literature on sex offender risk assessment could benefit researchers, reviewers, and readers alike.

Conflict of interest statement

Funding

References

1. Aebi, M., Plattner, B., Steinhausen, H.-C., & Bessler, C. (2011). Predicting sexual and nonsexual recidivism in a consecutive sample of juveniles convicted of sexual offences. Sexual Abuse: A Journal of Research and Treatment, 23(4), 456-473. doi:

10.1177/1079063210384634

2. Barnett, G. D., Wakeling, H. C., & Howard, P. D. (2010). An examination of the predictive validity of the Risk Matrix 2000 in England and Wales. Sexual Abuse: A Journal of Research and Treatment, 22(4), 443-470. doi: 10.1177/1079063210384274

3. Beggs, S. M., & Grace, R. C. (2010). Assessment of dynamic risk factors: An independent validation study of the Violence Risk Scale: Sexual offender version. Sexual Abuse: A Journal of Research and Treatment, 22(2), 234-251. doi: 10.1177/1079063210369014

4. Boer, D. P., Hart, S. D., Kropp, P. R., & Webster, C. D. (1997). Manual for the Sexual Violence Risk-20. Professional guidelines for assessing risk of sexual violence. Burnaby, BC: Simon Fraser University, Mental Health, Law, and Policy Institute.

5. Brouillette-Alarie, S., & Proulx, J. (2012). Predictive validity of the Static-99R and its dimensions. Journal of Sexual Aggression, 19(3), 311-328. doi: 10.1080/13552600.2012.747630

6. Brown, S., Harkins, L., & Beech, A. R. (2012). General and victim-specific empathy: Associations with actuarial risk, treatment outcome, and sexual recidivism. Sexual Abuse: A Journal of Research and Treatment, 24(5), 411-430. doi: 10.1177/1079063211423944

7. Caldwell, M. F. (2013). Accuracy of sexually violent person assessments of juveniles adjudicated for sexual offenses. Sexual Abuse: A Journal of Research and Treatment, 25, 516-526. doi:10.1177/1079063213480818

8. Chu, C. M., Ng, K., Fong, J., & Teoh, J. (2012). Assessing youth who sexually offended: The predictive validity of the ERASOR, J-SOAP-II, and YLS/CMI in a non- Western context. Sexual Abuse: A Journal of Research and Treatment, 24(2), 153-174. doi: 10.1177/1079063211404250

9. Craissati, J., Bierer, K., & South, R. (2010). Risk, reconviction and "sexually risky behaviour" in sex offenders. Journal of Sexual Aggression, 17(2), 153-165. doi: 10.1080/13552600.2010.490306

10. Duwe, G., & Freske, P. J. (2012). Using logistic regression modeling to predict sexual recidivism: The Minnesota Sex Offender Screening Tool-3 (MnSOST-3). Sexual Abuse: A Journal of Research and Treatment. doi: 10.1177/1079063211429470

11. Eher, R., Matthes, A., Schilling, F., Haubner-MacLean, T., & Rettenberger, M. (2012). Dynamic risk assessment in sexual offenders using STABLE-2000 and the STABLE-2007: An investigation of predictive and incremental validity. Sexual Abuse: A Journal of Research and Treatment, 24(1), 5-28. doi: 10.1177/1079063211403164

12. Fanniff, A. M., & Letourneau, E. J. (2012). Another piece of the puzzle: Psychometric properties of the J-SOAP-II. Sexual Abuse: A Journal of Research and Treatment, 24(4), 378-408. doi: 10.1177/1079063211431842

13. Fava, G. A. (2007). Financial conflicts of interest in psychiatry. World Psychiatry, 6, 19-24.

14. Fisher, R. A. (1922). On the interpretation of χ2 from contingency tables, and the calculation of P. Journal of the Royal Statistical Society, 85, 87-94. doi:10.2307/2340521

15. Grady, M. D., Edwards, D., Pettus-Davis, C., & Abramson, J. (2013). Does volunteering for sex offender treatment matter? Using propensity score Analysis to understand the effects of volunteerism and treatment on recidivism. Sexual Abuse: A Journal of Research and Treatment, 25(4), 319-346. doi: 10.1177/1079063212459085

16. Griffin, H. L., Beech, A., Print, B., Bradshaw, H., & Quayle, J. (2008). The development and initial testing of the AIM2 framework to assess risk and strengths in young people who sexually offend. Journal of Sexual Aggression, 14(3), 211-225. doi: 10.1080/13552600802366593

17. Griffin, H. L., & Vettor, S. (2011). Predicting sexual re-offending in a UK sample of adolescents with intellectual disabilities. Journal of Sexual Aggression, 18(1), 64-80. doi: 10.1080/13552600.2011.617013

18. Grubin, D. (2011). A large-scale evaluation of Risk Matrix 2000 in Scotland. Sexual Abuse: A Journal of Research and Treatment, 23(4), 419-433. doi: 10.1177/1079063211411309

19. Hall, G. C. N. (1990). Prediction of sexual aggression. Clinical Psychology Review, 10, 229-245. doi:10.1016/0272-7358(90)90059-J

20. Hanson, R. K., & Thornton, D. (2003). Notes on the development of Static-2002. Ottawa, ON: Department of the Solicitor General of Canada.

21. Hanson, R. K., & Thornton, D. (1999). Static-99: Improving actuarial risk assessments for sex offenders (User Report 99-02). Ottawa, ON: Department of the Solicitor General of Canada.

22. Harkins, L., Beech, A. R., & Goodwill, A. M. (2010). Examining the influence of denial, motivation, and risk on sexual recidivism. Sexual Abuse: A Journal of Research and Treatment, 22(1), 78-94. doi: 10.1177/1079063209358106

23. Helmus, L., Hanson, K., & Morton-Bourgon, K. E. (2011). International comparisons of the validity of actuarial risk tools for sexual offenders, with a focus on Static-99. In Boer, D. P., Eher, R., Craig, L. A., Miner, M. H. & Pfäfflin, F. (Eds.), International perspectives on the assessment and treatment of sexual offenders (pp. 55-83): John Wiley & Sons, Ltd.

24. Helmus, L., Hanson, R. K., Thornton, D., Babchishin, K. M., Harris, A. J. R. (2012). Absolute recidivism rates predicted by Static-99R and Static-2002R sex offender risk assessment tools vary across samples: A meta-analysis. Criminal Justice and Behavior, 33, 1148-1171. doi:10.1177/0093854812443648

25. Kingston, D. A., Yates, P. M., Firestone, P., Babchishin, K., & Bradford, J. M. (2008). Long-term predictive validity of the Risk Matrix 2000: A comparison with the Static-99 and the Sex Offender Risk Appraisal Guide. Sexual Abuse: A Journal of Research and Treatment, 20(4), 466-484. doi: 10.1177/1079063208325206

26. Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. Biometrics, 33, 159-174.

27. Larntz, K. (1978). Small-sample comparisons of exact levels for chi-squared goodness-of-fit statistics. Journal of the American Statistical Association, 73, 253-263. doi:10.2307/2286650

28. Maj, M. (2005). Conflicts of interests in psychiatric practice and research. Psychiatrie, 3, 138- 140.

29. McGrath, R. J., Lasher, M. P., & Cumming, G. F. (2012). The Sex Offender Treatment Intervention and Progress Scale (SOTIPS): Psychometric properties and incremental predictive validity with Static-99R. Sexual Abuse: A Journal of Research and Treatment, 24(5), 431-458. doi: 10.1177/1079063211432475

30. Montana, S., Thompson, G., Ellsworth, P., Lagan, H., Helmus, L., & Rhoades, C. J. (2012). Predicting relapse for catholic clergy sex offenders: The use of the Static-99. Sexual Abuse: A Journal of Research and Treatment, 24(6), 575-590. doi: 10.1177/1079063212445570

31. Olver, M. E., Nicholaichuk, T. P., Gu, D., & Wong, S. C. P. (2013). Sex offender treatment outcome, actuarial risk, and the aging sex offender in Canadian corrections: A long-term follow-up. Sexual Abuse: A Journal of Research and Treatment, 25(4), 396-422. doi: 10.1177/1079063212464399

32. 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.

33. Rossegger, A., Gerth, J., Urbaniok, F., Laubacher, A., & Endrass, J. (2010). Der Sex Offender Risk Appraisal Guide (SORAG) - Validität und autorisierte deutsche Übersetzung. [The Sex Offender Risk Appraisal Guide (SORAG) - validity and authorized German translation]. Fortschr Neurol Psychiatr, 78, 658-667. doi: 10.1055/s-0029-1245688

34. Rutjes, A. W. S., Reitsma, J. B., Di Nisio, M., Smidt, N., van Rijn, J. C., & Bossuyt, P. M. M. (2006). Evidence of bias and variation in diagnostic accuracy studies. Canadian Medical Association Journal, 174, 469-476.

35. Singh, J. P., Desmarais, S., & Van Dorn, R. A. (2013). Measurement of predictive validity in violence risk assessment studies: A second-order systematic review. Behavioral Sciences and the Law, 31, 55-73.
36. Singh, J. P., Fazel, S., Gueorguieva, R., & Buchanan, A. (2013). Rates of sexual recidivism in high risk sex offenders: A meta-analysis of 10,422 participants. Sexual Offender Treatment, 7, 44-57.
37. Singh, J. P., Grann, M., & Fazel, S. (2013). Authorship bias in violence risk assessment? A systematic review and meta-analysis. PLoS ONE, 8, e72484.
38. Singh, J. P., Grann, M., & Fazel, S. (2011). A comparative study of risk assessment tools: A systematic review and metaregression analysis of 68 studies involving 25,980 participants. Clinical Psychology Review, 31, 499-513. doi: 10.1016/j.cpr.2010.11.009
39. Singh, J. P., Yang, S., Mulvey, E., & the RAGEE Group. (in press). Reporting standards for risk assessment predictive validity studies: The Risk Assessment Guidelines for the Evaluation of Efficacy (RAGEE) Statement. Law and Human Behavior.
40. Skelton, A., & Vess, J. (2008). Risk of sexual recidivism as a function of age and actuarial risk. Journal of Sexual Aggression, 14(3), 199-209. doi: 10.1080/13552600802267098
41. Smallbone, S., & Rallings, M. (2013). Short-term predictive validity of the Static-99 and Static-99-R for indigenous and nonindigenous Australian sexual offenders. Sexual Abuse: A Journal of Research and Treatment, 25(3), 302-316. doi: 10.1177/1079063212472937
42. Storey, J. E., Watt, K. A., Jackson, K. J., & Hart, S. D. (2012). Utilization and implications of the Static-99 in practice. Sexual Abuse: A Journal of Research and Treatment, 24(3), 289-302. doi: 10.1177/1079063211423943
43. Swinburne Romine, R. E., Miner, M. H., Poulin, D., Dwyer, S. M., & Berg, D. (2012). Predicting reoffense for community-based sexual offenders: An analysis of 30 years of data. Sexual Abuse: A Journal of Research and Treatment, 24(5), 501-514. doi: 10.1177/1079063212446514
44. Thornton, D., Mann, R. E., Webster, S., Blud, L., Travers, R., Friendship, C. et al. (2003). Distinguishing and combining risks for sexual and violent recidivism. In R. A. Prentky, E. S. Janus, & M. C. Seto (Eds), Sexually coercive behavior: Understanding and management (pp. 225-235). Annals of the New York Academy of Science, Vol. 989. New York: New York Academy of Sciences.
45. Wakeling, H. C., Howard, P., & Barnett, G. (2011). Comparing the validity of the RM2000 scales and OGRS3 for predicting recidivism by internet sexual offenders. Sexual Abuse: A Journal of Research and Treatment, 23(1), 146-168. doi: 10.1177/1079063210375974
46. Wilcox, D., Beech, A., Markall, H. F., & Blacker, J. (2009). Actuarial risk assessment and recidivism in a sample of UK intellectually disabled sexual offenders. Journal of Sexual Aggression, 15(1), 97-106. doi: 10.1080/13552600802578577
47. Wilson, R. J., Abracen, J., Looman, J., Picheca, J. E., & Ferguson, M. (2011). Pedophilia: An evaluation of diagnostic and risk prediction methods. Sexual Abuse: A Journal of Research and Treatment, 23(2), 260-274. doi: 10.1177/1079063210384277
48. Worling, J. R., Bookalam, D., & Litteljohn, A. (2012). Prospective validity of the Estimate of Risk of Adolescent Sexual Offense Recidivism (ERASOR). Sexual Abuse: A Journal of Research and Treatment, 24(3), 203-223. doi: 10.1177/1079063211407080
49. Zucker, K. J. (2013). The impact factor: Just the facts. Archives of Sexual Behavior. doi:10.1007/s10508-013-0105-1

Table 1:
Reporting quality of two leading sex offender risk assessment journals according to Risk Assessment Guidelines for the Evaluation of Efficacy (RAGEE) Statement checklist criteria

RAGEE Checklist Section	Item	Description	SAJRT (k, %)	JSA (k, %)	chi²
ABSTRACT	R1	Include a structured abstract describing the study	0 (0)	0 (0)	
	R2	Identify the article as a risk assessment study in which predictive validity is measured	20 (95.2)	6 (100.0)	
	R3	Identify the risk assessment instrument(s) whose predictive validity is measured	21 (100.0)	4 (66.7)	*
	R4	State the nature of the principal outcome (e.g., violence, sexual violence, criminal offending, institutional misconduct)	17 (81.0)	6 (100.0)	
INTRODUCTION	R5	Provide the rationale and a summary of the scientific/theoretical background for the study	21 (100.0)	6 (100.0)	
	R6	State the research questions and/or study aims	21 (100.0)	6 (100.0)	
METHODS					
Participants	R7	Report the sample size	21 (100.0)	6 (100.0)	
	R8	Report the sex/gender composition of the sample	17 (81.0)	4 (66.7)	
	R9	Report the average age at assessment (with dispersion parameter)	1 (4.8)	0 (0)	
	R10	Report the index offense composition of the sample	7 (33.3)	2 (33.3)	
	R11	Report the characteristics of groups that underwent subgroup analysis	20 (95.2)	6 (100.0)	
Instrument Design	R12	Report the acronym(s) and full name(s) of the instrument(s) under investigation with appropriate reference to source document	21 (100.0)	4 (66.7)	*
	R13	Report the number of items on the instrument(s) under investigation	21 (100.0)	5 (83.3)	
	R14	Report the approach by which the assessment information from the instrument(s) under investigation is organized into an overall evaluation of risk	19 (90.5)	6 (100.0)	
	R15	Report the population for which the instrument(s) under investigation was intended to be used	21 (100.0)	4 (66.7)	*
	R16		9 (42.9)	0 (0)	

		Report the outcome(s) that the instrument(s) under investigation was intended to assess		
	R17	Report the length of follow-up for which manual-recommended probability estimates of risk were derived for the instrument(s) under investigation	8 (38.1)	3 (50.0)
	R18	Report the cut-off score(s) and/or risk categories that the instrument(s) under investigation was designed to use to classify risk level	16 (76.2)	5 (83.3)
Instrument Administration	R19	Report whether risk assessments were conducted in the context of research or practice	20 (95.2)	6 (100.0)
	R20	Identify when risk assessments occurred (e.g., pre-admission, admission, release, post-release)	8 (38.1)	2 (33.3)
	R21	Report the number of assessors in the study as well as their training in the administration of the instrument(s) under investigation	7 (33.3)	1 (16.7)
	R22	Identify the source(s) of information used to administer the instrument(s) under investigation	18 (85.7)	6 (100.0)
	R23	Describe any modifications made to the instrument(s) under investigation	21 (100.0)	6 (100.0)
Study Design	R24	Report the geographical location and clinical setting in which risk was assessed	20 (95.2)	6 (100.0)
	R25	Describe the method(s) used to recruit participants	11 (52.4)	4 (66.7)
	R26	Identify the temporal design of the study (prospective or quasi-prospective)	14 (66.7)	5 (83.3)
	R27	Identify the setting in which participants were followed to ascertain whether the outcome(s) of interest had occurred	19 (90.5)	3 (50.0)
	R28	Report the average length of follow-up and time at risk (with dispersion parameter, if not fixed), including a description of periods subtracted from follow-up time (e.g., incarceration and/or hospitalization)	21 (100.0)	6 (100.0)
Predicted Outcome	R29	Specify the event(s) coded as meeting outcome criteria (e.g., assault, rape, homicide)	9 (42.9)	0 (0)
	R30	Identify the type (e.g., arrest, charge, conviction, incarceration) and source (e.g., criminal records, self-report, collateral) of information used to detect outcome occurrence	21 (100.0)	4 (66.7) *

Statistical Analysis	R31	Describe the statistical methods used to conduct all analyses, and report the purpose of each analysis	18 (85.7)	4 (66.7)	
	R32	Report whether risk scores and/or risk categories of the instrument(s) under investigation were used as an independent variable in analyses	21 (100.0)	6 (100.0)	
	R33	Identify the statistical significance level used	15 (71.4)	1 (16.7)	*
	R34	Describe any subgroup analyses planned a priori	21 (100.0)	5 (83.3)	
	R35	Report inter-rater reliability for administration of the instrument(s) under investigation (if conducted). If inter-rater reliability was not assessed, clarify why not	10 (47.6)	1 (16.7)	
	R36	Include at least one discrimination performance indicator when measuring predictive validity	21 (100.0)	6 (100.0)	
RESULTS					
Participant Outcomes	R37	Report the rate of attrition	21 (100.0)	6 (100.0)	
	R38	Report the outcome occurrence rate for the entire sample as well as for relevant subgroups	20 (95.2)	5 (83.3)	
Predictive Validity	R39	Report predictive validity performance indicators for each outcome of interest as specified in the Methods with associated dispersion parameters	21 (100.0)	6 (100.0)	
	R40	Report the number of participants with each risk score and/or in each risk category and how many went on to engage in the outcome(s) of interest	9 (42.9)	3 (50.0)	
	R41	Report the results of subgroup analyses planned a priori as specified in the Methods	21 (100.0)	6 (100.0)	
	R42	Describe and report the findings of any post hoc analyses conducted	21 (100.0)	6 (100.0)	
DISCUSSION	R43	Provide a summary of the principal findings, including a discussion of their relevance in the context of the current literature	21 (100.0)	6 (100.0)	
	R44	Discuss limitations of the study design	17 (81.0)	6 (100.0)	
	R45	Discuss the generalizability of study findings	17 (81.0)	2 (33.3)	*
	R46	Discuss future research directions based on study findings	19 (90.5)	6 (100.0)	

DISCLOSURES	R47	Report any commercial interests and/or source(s) of funding as well as their role(s) in the conduct of the study	21 (100.0)	5 (83.3)	
	R48	Report whether an author or translator of the risk assessment instrument(s) under investigation was also a study author	21 (100.0)	6 (100.0)	
	R49	Report whether the study presented in the article has been published in an alternative form (e.g., government report)	21 (100.0)	6 (100.0)	
	R50	Report whether the sample or a portion thereof has been studied in other publications	21 (100.0)	6 (100.0)	

Note. *$p < 0.05$; k = number of articles; SAJRT = Sexual Abuse: A Journal of Research and Treatment (k total = 21); JSA = Journal of Sexual Aggression (k total = 6); chi^2 = Fisher's exact test.

Supplement 1: Sex- and gender-related journals identified in systematic search process

- Archives of Sexual Behavior
- AIDS
- AIDS and Behavior
- AIDS Care
- AIDS Education and Prevention
- AIDS Patient Care and STDs
- AIDS Reviews
- Asian Journal of Andrology
- Asian Journal of Women s Studies
- Australian Feminist Studies
- Body & Society
- Body Image
- Culture, Health, and Sexuality
- Differences: A Journal of Feminist Cultural Studies
- Empirical Journal of Same-Sex Sexual Behavior
- European Journal of Women s Studies
- Evolution and Human Behavior
- Evolutionary Psychology
- Feminism and Psychology
- Feminist Review
- Feminist Studies
- Feminist Theory
- Feministische Studien
- Fertility and Sterility
- Frontiers: Journal of Women s Studies
- Gender and Education
- Gender and Society
- Gender, Place, and Culture
- GLQ: A Journal of Gay and Lesbian Studies

- Hormone Research in Pediatrics
- Hormones and Behavior
- Human Reproduction
- Indian Journal of Gender Studies
- International Journal of Andrology
- International Journal of Impotence Research
- International Journal of Sexual Health
- International Journal of STD and AIDS
- International Perspectives on Sexual and Reproductive Health
- Journal of Acquired Immune Deficiency Syndromes
- Journal of Andrology
- Journal of Child Abuse & the Law
- Journal of Child Sexual Abuse
- Journal of Gender Studies
- Journal of the History of Sexuality
- Journal of Homosexuality
- Journal of Men s Health
- Journal of Psychosomatic Obstetrics and Gynecology
- Journal of Sex & Marital Therapy
- Journal of Sex Research
- Journal of Sexual Aggression
- Journal of Sexual Medicine
- Journal of Sexual Offender Civil Commitment: Science and the Law
- Journal of Women s Health
- Men and Masculinities
- Menopause
- Paidika: The Journal of Paedophilia
- Perspectives on Sexual and Reproductive Health
- Politics and Gender
- Psychology of Men & Masculinity
- Psychology of Women Quarterly
- Psychoneuroendocrinology
- Sex Roles
- Sexual Abuse: A Journal of Research and Treatment
- Sexual and Relationship Therapy
- Sexual Development
- Sexual Health
- Sexual Offender Treatment
- Sexualities
- Sexuality and Disability
- Sexuality Research and Social Policy
- Sexually Transmitted Diseases
- Sexually Transmitted Infections
- Signs
- Women & Health
- Women and Therapy
- Women s Health Issues
- Women s Studies International Forum
- Zeitschrift fur Sexualforschung

Author address

Jay P. Singh, PhD
Global Institute of Forensic Research
11790 Plaza America Drive, Suite 810
Reston, VA 20190, United States of America
jaysingh@gifrinc.com