

Assessment of Relative Risk for Sexual and Violent Recidivism With Risk Matrix 2000

David Thornton^a, Sabrina Eberhaut^b, Gina Ambroziak^c, Reinhard Eher^b

[a] FAsTR LLC, Madison, WI, USA. [b] Federal Evaluation Centre for Violent and Sexual Offenders (FECVSO), Ministry of Justice, Vienna, Austria. [c] Sand Ridge Secure Treatment Center, Mauston, WI, USA.

Sexual Offending: Theory, Research, and Prevention, 2020, Vol. 15(1), Article e3129,
<https://doi.org/10.5964/sotrap.3129>

Received: 2020-03-21 • Accepted: 2020-04-10 • Published (VoR): 2020-07-22

Handling Editor: Martin Rettenberger, Centre for Criminology (Kriminologische Zentralstelle – KrimZ), Wiesbaden, Germany

Corresponding Author: David Thornton, FAsTR LLC, 1213 N Sherman Avenue Suite 334, Madison, WI 53704, United States. E-mail: davidsmthornton@icloud.com

Abstract

Risk Matrix 2000's ability to assess relative risk for sexual and violent recidivism has been well established through meta-analysis. However, the instrument was originally designed for use in the United Kingdom and has not been widely tested in other parts of Europe, raising questions about how generalizable the results are. This paper assessed the instrument's ability to assess relative risk for these outcomes in a sample of over 300 Austrian adult males serving a prison sentence for a sexual offense for whom 5-year rates of sexual and violent recidivism were available. Results indicated an ability to assess relative risk that was comparable to that observed in the United Kingdom. In the context of previous results from Germany, the findings of the present study indicate that Risk Matrix 2000 may be appropriately applied in Europe. Analyses explored the added value of using the Sexual and Violence risk scales in conjunction to identify the kinds of criminogenic need that should be addressed in treatment and supervision.

Keywords

sexual recidivism, actuarial risk assessment, RM2000



Highlights

- The Risk Matrix 2000 S-scale assessed relative risk for sexual recidivism as well in Austria as it has in the United Kingdom where it was developed.
- The Risk Matrix 2000 C-scale assessed relative risk for violence (including contact sex offending) as well in Austria as it has in the United Kingdom where it was developed.
- Considered together, the Risk Matrix S and V scales can suggest whether individuals with a history of sexual offending should be prioritized for sex offense specific treatment or whether targeting more general criminogenic needs should be prioritized.
- Taken with previous research in Germany, this suggests that Risk Matrix 2000 can appropriately be used to assess relative risk for sexual and violent recidivism in Europe.

Professionals are commonly asked to assess the risk presented by men who have been convicted of committing sexual offenses. Professional guidelines (e.g., [ATSA, 2014](#)) emphasize the need to use structured empirically-based instruments in the assessment of risk. More specifically, for the assessment of sexual recidivism risk, meta-analytic comparison suggests that empirical-actuarial instruments are more predictive than structured professional judgment ([Hanson & Morton-Bourgon, 2009](#)).

Although these assessments typically emphasize risk for sexual recidivism, men convicted of sexual offenses typically pose about equal risks for sexual and non-sexual violent recidivism ([Rettenberger et al., 2015](#); [Thornton et al., 2003](#)). This suggests that a comprehensive assessment of the risks for serious negative outcomes for these men should take into account the potential for both sexually and non-sexually violent offenses.

Risk Matrix 2000 (RM2000) is a static actuarial instrument designed to assist professionals in assessing the risk posed by adult males who have been convicted for a sexual offense ([Thornton et al., 2003](#)). It is somewhat similar to the widely used actuarial, Static-99 ([Hanson & Thornton, 2000](#)), but is easier to score and so potentially can be used by a wider range of professionals or when less information is available. It makes finer distinctions for age at release than does Static-99 but less than the revised version, Static-99R ([Helmus et al., 2012](#)).

RM2000 allows estimating the risk for sexual recidivism (S-scale) and for nonsexual violence (V-scale), and the combination of these two outcomes (C-scale), to be assessed in a separate but integrated way, with scales that have been co-normed on the same samples with the results being expressed in similar relative risk categories for both outcomes. This is something that is not currently possible for Static-99R although the STATIC family of instruments now includes a scale to assess risk for non-sexual violence

(the BARR-2002R: Babchishin et al., 2016) that can be used as part of an assessment with Static-2002R.

The predictive accuracy of RM2000 was examined by Helmus et al. (2013) in a meta-analysis of 16 distinct samples. This found an ability to assess relative risk that compared well to that found for other instruments designed to assess men convicted for sexual offenses. However, RM2000 was developed in the United Kingdom (UK), and 10 of the 16 samples were from the UK, as were over 70% of the participants. Only one of the samples was from continental Europe. Furthermore, these authors reported that the predictive accuracy of RM2000 was higher in samples from the UK than in samples from other countries, raising questions about how far results might be generalizable across jurisdictions and emphasizing a need for further study of RM2000's predictive properties in non-UK samples from Europe.

Lehmann et al. (2016) reported the development of international recidivism norms for RM2000 based on four "Routine" samples, drawn from England and Wales, Scotland, Germany, and Canada. "Routine" here refers to samples that are reasonably representative of persons convicted for sexual offenses and sentenced to either community supervision or imprisonment. Similar levels of accuracy in assessing relative risk were observed across these four Routine samples. This suggests that the previously observed differences between the UK and other jurisdictions in the accuracy with which relative risk is assessed may reflect the type of sample rather than differences in jurisdictions.

Purpose of the Present Study

The primary aim of the current study was to evaluate how well relative risk for sexual and violent recidivism was assessed by RM2000 in a sample of adult males released from Austrian prisons following a sentence imposed for a sexual offense. More specifically it was intended to:

1. Assess the accuracy with which relative risk for sexual recidivism was assessed by RM2000's S scale.
2. Assess the accuracy with which relative risk for violent recidivism (defined here as new convictions for contact sex offenses or non-sexual violence) was assessed by RM2000's C scale.
3. Test the hypothesis that both the S and V scales would make independent contributions to the prediction of violent recidivism.
4. Test the hypothesis that the S scales would make an independent contribution to the prediction of sexual recidivism while the V scale would not.
5. Explore similarities and differences in the patterns of recidivism for individuals who rated at least High on the S-scale alone, the V-scale alone, or both.

Method

Participants

The study sample was comprised of $N = 339$ male sexual offenders who had been sentenced for a prison term. The mean age of the sample at the time of admission to the Federal Evaluation Centre for Violent and Sexual Offenders (FECVSO) was 40.26 years (range 17 to 70 years, 95% CI bootstrap [38.9, 41.6]). Just under half of the participants (45.4%) had sexually offended against adults, whereas just over half (54.6%) had offended against minors below age of 14. Of the total sample, $N = 337$ could be followed up after prison release. Mean age at time at release was 42.4 years (range 17.7 to 71.8 years, 95% CI bootstrap [41.1, 43.8]). Follow-up time was 10.19 years on average (range 3.8 to 12.8 years, 95% CI bootstrap [10.1, 10.3]).

Measures

Risk Matrix 2000 (RM2000)

As described earlier, RM2000 is a static actuarial risk assessment instrument designed to assess the risk presented by adult males convicted of a sexual offense (Thornton et al., 2003). It comprises three scales. The S-scale assesses risk for sexual recidivism, the V-scale the risk for nonsexual violence, and the C-scale the risk for violent recidivism (defined as either recidivism that involves either sexual offenses or non-sexual violence). The item content is described in Thornton et al. (2003), and in Helmus et al. (2013) and a form describing how the items are integrated is in the Appendix. The three scales produce categories of relative risk (Low, Moderate, High, Very High) rather than scores. These correspond to Below Average (I and II), Average, Above Average, and Well Above Average from the standardized risk assessment categories (Hanson et al., 2017). Wakeling et al. (2011) report inter-rater reliability (ICC) of .90 and above. Helmus et al. (2013) reported average d -statistics of 0.74 for the S-scale predicting sexual recidivism and 0.81 for the C-scale predicting violent recidivism. RM2000 scores were available for 339 cases.

Risk Matrix 2000 was implemented in the present research by having the items included into a spreadsheet that was being used to code data for a larger project. The Risk Matrix scoring guide was used to write the instructions for coding the items. The items were coded by the second author following training by the first author who was one of the creators of Risk Matrix 2000. The spreadsheet had been programmed by the third author so that it automatically created Risk Matrix classifications once item ratings had been input. At the time the second author was working under the supervision of the fourth author.

Recidivism

Sexual recidivism was defined as reconviction for a sexual offense. Violent recidivism was defined as reconviction for a contact sexual offense or for a non-sexual violent offense. Information about recidivism was obtained from the Austrian Ministry of Internal Affairs.

Measuring Relative Risk

Accuracy in assessing relative risk can be assessed with statistics like the AUC or the standardized mean difference between recidivists and non-recidivists (Cohen's *d*: Cohen, 1988). It may also be assessed through fitting a logistic regression equation predicting recidivism from a scale score with the b_1 (slope) coefficient if a fixed follow up is used. Alternatively, if a variable follow up is used with hazard ratios analyzed using Cox Regression then Harrell's C statistic (Harrell et al., 1982) is analogous to the AUC and can be interpreted as the probability that, given two randomly selected individuals, the one with the higher score will reoffend first. Five-year follow up was available for 332 of the 339 cases while some follow up (including for less than five years) was available for 337 cases.

Compliance With Research Ethics

It is the inherent obligation of the Federal Evaluation Centre for Violent and Sexual Offenders (FECVSO) to continuously evaluate the accuracy of its risk assessment approaches. The FECVSO is a subdivision of the Austrian Ministry of Justice. This evaluation study, therefore, was performed in line with the legal and ethical standards of the Austrian Ministry of Justice and the National and European Data Protection Act. It also relates to the Recommendations of the Committee of Ministers to Member States (Council of Europe) Concerning Dangerous Offenders and the Directives of the European Parliament and of the Council on Combating the Sexual Abuse and Sexual Exploitation of Children. All high ethical and legal standards concerning research on a vulnerable population have been adhered to accordingly.

Results

For the S-scale for the full sample ($N = 339$), 37.5% fell in the Low category, 38.3% fell in the Moderate category, 16.2% fell in the High category, and 8.0% fell in the Very High category. For the C-scale, 29.5% fell in the Low category, 33.3% fell in the Moderate category, 30.7% fell in the High category, and 6.5% fell in the Very High category.

Forty participants (11.9%) reoffended sexually over the total follow up period, and ninety-one (27.0%) relapsed with a violent offense during follow-up time. Five-year sexual and violent recidivism rates were 6.9% (23 of 332) and 19.3% (64 of 332) respectively.

The AUC for the S-scale assessing risk for sexual recidivism was .78 (95% CI [.69, .87]) while the AUC for the C-scale assessing risk for violent recidivism was .73 (95% CI [.67, .80]). The odds ratios ($\text{Exp}(b)$) from the corresponding logistic regression equations were 3.09 (95% CI [1.95, 4.90]) for the S-scale predicting sexual recidivism and 2.75 (95% CI [1.94, 3.89]) for the C-scale predicting violent recidivism.

When a logistic regression equation predicting sexual recidivism from both the S and V scales was fitted the regression weight for the S-scale was statistically significant ($b = +1.16$; $p < .001$; $\text{Exp}(b) = 3.20$, 95% CI [1.97, 5.23]) while the regression weight for the V-scale was not ($b = -0.11$; $p = .664$; $\text{Exp}(b) = 0.89$, 95% CI [0.54, 1.49]). In contrast, when a logistic regression equation predicting violent recidivism from both the S and V scales was fitted the regression weight for the S-scale was statistically significant ($b = +0.41$; $p = .020$; $\text{Exp}(b) = 1.51$, 95% CI [1.07, 2.14]) while the regression weight for the V-scale was also significant ($b = +0.70$; $p < .001$; $\text{Exp}(b) = 2.01$, 95% CI [1.43, 2.81]).

Over half of the cases (199) fall in the moderate or low categories on both S and V scales. This group had a five-year sexual recidivism rate of 2.5% and a five-year violent recidivism rate of 9.5%. Some 43 cases fell in the moderate or low categories on the V scale but in the High or Very High categories on the S scale. This group had a five-year sexual recidivism rate of 18.6% and a five-year violent recidivism rate of 25.6%. Some 55 cases fell in the moderate or low categories on the S scale but in the High or Very High categories on the V scale. This group had a five-year sexual recidivism rate of 7.3% and a five-year violent recidivism rate of 32.7%. Some 35 cases fell in the High or Very High categories on both S and V scales. This group had a five-year sexual recidivism rate of 17.1% and a five-year violent recidivism rate of 45.7%.

Discussion

The aim of this paper was to investigate the ability of Risk Matrix 2000 (RM2000) scales to assess relative risk for sexual and violent recidivism in a sample from outside the UK. Previous research had suggested that the ability to assess relative risk was greater in the UK where RM2000 was developed than in other countries (Helmus et al., 2013). The present results do not fit that generalization, at least for the S-scale. In the present sample the S-scale's ability to assess relative risk for sexual recidivism was greater than that typically reported in other samples. For example, odds ratios in the four samples analyzed by Lehmann et al. (2016) ranged from a low of 2.28 to a high of 2.74, while that observed here was 3.20. AUCs for the two UK cross-validation samples for the S-scale in Thornton et al. (2003) were .77 and .75, while that observed here was .78. Regarding the C-scale, the odds ratios in the four Lehmann et al. samples ranged from 2.43 to 2.83, while it was 2.75 in the present sample. AUCs for the two UK cross-validation samples for the C-scale in Thornton et al. (2003) were .81 and .74, while that observed here was .73.

In short, the ability of the RM2000 scales to assess relative risk for sexual and for violent recidivism in the present sample was similar to or better than that observed in other samples, including samples from the UK. This suggests that the previously observed variation in ability to assess relative risk may have had more to do with sample type and how recidivism was measured than with jurisdiction.

Turning to recidivism rates reported for groups produced by cross-classifying individuals by whether they fell into the higher risk categories (High and Very High) on the S and V scales, it is useful to consider the potential implications for risk management services for the different groups. First, however, it is important to acknowledge that neither five-year sexual recidivism rates nor five-year violent recidivism rates were high in absolute terms. This reinforces the importance of explaining that the category labels are intended to communicate relative rather than absolute risk.

For the large group that did not fall into the higher risk categories on either scale, the observed sexual recidivism rates do not suggest that allocating significant specialized sexual-offense-specific treatment services is warranted. Indeed, the risk they do present is more for non-sexual violence as their violent recidivism rate is about four times their sexual recidivism rate. Turning to those who fall in the higher risk categories only on the V scale, the risk they present for violence is very largely comprised of non-sexual violence, with the violent recidivism rate being more than four times the sexual recidivism rate. Thus, despite their history of sexual offending, the criminogenic needs associated with non-sexual violence should be prioritized for treatment and management purposes.

In contrast, those in the higher categories only on the S scale present a risk for violence that is very largely comprised of risk for sexual recidivism. Nearly three-quarters of their violent recidivism involves sexual offending. This is the only group for whom a largely exclusive focus on criminogenic needs associated with sexual recidivism seems warranted. Finally, those rated as being in the higher risk categories on both scales, clearly warrant sex offense specific treatment services; however, the services they receive also need to be oriented to the prevention of non-sexual violence since at least half of their risk for violence is for non-sexual violence. This last group is the one for whom intensive services seem most warranted as their absolute reconviction rate approaches 50% over five years and presumably would be much higher with an extended follow up. This group represents only about a tenth of the sample.

Conclusions

Three practical implications are suggested by the above discussion. First, it is safe to use RM2000 to assess relative risk in Austria. Furthermore, given that similarly good results were reported in Germany by [Lehmann et al. \(2016\)](#) it would seem reasonable to employ RM2000 in other European countries. Second, there is likely to be clinical value in using the S and V scales together to profile individuals being assessed. The combined results have implications for treatment and risk management that go beyond what is to

be obtained with the S scale alone. Use of the V scale can also serve as an important reminder that the risk associated with a history of sexual offending is not solely, or even necessarily primarily, the risk of further sexual offending. Third, only a small proportion of those imprisoned for a sexual offense present a level of risk that warrants intensive treatment services. Correctional systems should use systematic assessment to focus treatment and management resources on this group.

Limitations and Future Directions

The primary limitation of the present study is the sample size. This, especially for the higher risk categories, precludes examination of the degree to which absolute risk estimates correspond to those in the International Norms. More differentiated analyses would also have been possible if non-sexual violence had been coded as a separate outcome rather than folded into a more general violence category that also included sex offending.

A second limitation is that the outcome variables available for analysis do not correspond exactly to those for which Risk Matrix 2000 was designed. Specifically, here Violent recidivism was defined as new conviction for a contact sexual offense or for a nonsexual violent offense. In contrast, in the original Risk Matrix research violent recidivism referred solely to recidivism involving a nonsexual violent offense and it was this outcome that the V-scale was designed to predict. Because of the way the outcome data was collected this variable was, however, not available to us. Similarly, in the original Risk Matrix research a combined sexual plus other violent recidivism variable was used, and it was that outcome that the C-scale was designed to predict. This outcome variable too was not available to us although the Violent recidivism variable that was available was close to it. The main consequence of these limitations was that we were unable to test the V-scale against the outcome it was designed to predict (non-sexual violence).

The present work might be developed in two ways. First, similar studies carried out in other European countries would further explore the generalizability of the results, and, if integrated through meta-analysis, might allow a more precise analysis of the behavior of individuals falling in the relatively rare higher risk categories. This would also allow the calibration of the estimates of absolute risk from [Lehmann et al. \(2016\)](#) to be evaluated. Second, the degree to which psychological risk factors can add to assessment of relative risk could be explored. It is notable that in previous studies of this Austrian cohort, the added value of psychological assessment has sometimes been limited (e.g. [Eher et al., 2012](#)).

Funding: The authors have no funding to report.

Competing Interests: The fourth author (RE) is one of SOTRAP's Editors-in-Chief, but played no editorial role for this particular article or intervened in any form in the peer review procedure.

Acknowledgments: The authors have no support to report.

Data Availability: The data used in this study are the property of the Austrian government and the authors do not have permission to post the underlying data files to a public website or to share it with other researchers.

References

- Association for the Treatment of Sexual Abusers. (2014). *ATSA Practice Guidelines for the Assessment, Treatment, and Management of Male Adult Sexual Abusers*. Retrieved from http://www.atsa.com/Members/Adult/ATSA_2014_Adult_Practice_Guidelines.pdf
- Babchishin, K. M., Hanson, R. K., & Blais, J. (2016). Less is more: Using Static-2002R subscales to predict violent and general recidivism among sexual offenders. *Sexual Abuse, 28*(3), 187-217. <https://doi.org/10.1177/1079063215569544>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ, USA: Erlbaum.
- 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, 24*(1), 5-28. <https://doi.org/10.1177/1079063211403164>
- Hanson, R. K., Babchishin, K. M., Helmus, L. M., Thornton, D., & Phenix, A. (2017). Communicating the results of criterion referenced prediction measures: Risk categories for the Static-99R and Static-2002R sexual offender risk assessment tools. *Psychological Assessment, 29*(5), 582-597. <https://doi.org/10.1037/pas0000371>
- Hanson, R. K., & Morton-Bourgon, K. E. (2009). The accuracy of recidivism risk assessments for sexual offenders: A meta-analysis of 118 prediction studies. *Psychological Assessment, 21*(1), 1-21. <https://doi.org/10.1037/a0014421>
- Hanson, R. K., & Thornton, D. (2000). Improving risk assessments for sex offenders: A comparison of three actuarial scales. *Law and Human Behavior, 24*(1), 119-136. <https://doi.org/10.1023/A:1005482921333>
- Harrell, F. E., Jr., Califf, R. M., Pryor, D. B., Lee, K. L., & Rosati, R. A. (1982). Evaluating the yield of medical tests. *Journal of the American Medical Association, 247*(18), 2543-2546. <https://doi.org/10.1001/jama.1982.03320430047030>

- Helmus, L., Babchishin, K. M., & Hanson, R. K. (2013). The predictive accuracy of the Risk Matrix 2000: A meta-analysis. *Sexual Offender Treatment, 8*(2), 1-24.
<http://sexual-offender-treatment.org/125.html>
- Helmus, L., Thornton, D., Hanson, R. K., & Babchishin, K. M. (2012). Improving the predictive accuracy of Static-99 and Static-2002 with older sex offenders: Revised age weights. *Sexual Abuse: A Journal of Research and Treatment, 24*(1), 64-101.
<https://doi.org/10.1177/1079063211409951>
- Lehmann, R. J. B., Thornton, D., Helmus, L. M., & Hanson, R. K. (2016). Developing non-arbitrary metrics for risk communication: Norms for the Risk Matrix 2000. *Criminal Justice and Behavior, 43*(12), 1661-1687. <https://doi.org/10.1177/0093854816651656>
- Rettenberger, M., Briken, P., Turner, D., & Eher, R. (2015). Sexual offender recidivism among a population-based prison sample. *International Journal of Offender Therapy and Comparative Criminology, 59*(4), 424-444. <https://doi.org/10.1177/0306624X13516732>
- Thornton, D., Mann, R., Webster, S., Blud, L., Travers, R., Friendship, C., & Erikson, M. (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* [Annals of the New York Academy of Sciences, Vol. 989] (pp. 225-246).
<https://doi.org/10.1111/j.1749-6632.2003.tb07308.x>
- Wakeling, H. C., Mann, R. E., & Milner, R. J. (2011). Interrater reliability of Risk Matrix 2000/s. *International Journal of Offender Therapy and Comparative Criminology, 55*(8), 1324-1337.
<https://doi.org/10.1177/0306624X11386933>

Appendix: Risk Matrix 2000 Scoring Form

RM2000/S Scale – Risk for Sexual Recidivism

Step One: Scoring Risk Factors

Circle the number of points that apply for each risk factor.

Age	18-24 = 2 points; 25-34 = 1 point; Older = 0 points
Sexual Appearances	1 = 0 points; 2 = 1 point; 3,4 = 2 points; 5+ = 3 points
Criminal Appearances	4 or less = 0 points; 5 or more = 1 point

Step One: Categorization

Circle the total number of points from the previous table, and the corresponding Risk label.

Points	Risk
0	Low
1-2	Medium
3-4	High
5-6	Very High

Step Two: Aggravating Factors

Circle the number of points that apply for each aggravating factor.

Male Victim of Sex Offense	No = 0 points; Yes = 1 point
Stranger Victim of Sex Offense	No = 0 points; Yes = 1 point
Single (Never in Marital Type Relationship)	No = 0 points; Yes = 1 point
Non- Contact Sex Offence	No = 0 points; Yes = 1 point

Step Two: Revised Risk Category

Put the risk category up one (e.g. from I to II or from II to III, or from III to IV) if two or three aggravating factors apply, and up two categories (e.g. from I to III, or from II to IV) if four aggravating factors apply. Circle the Revised Risk Category and Label.

Revised Risk Category				
Label	Low	Medium	High	Very High

RM2000/V – Risk for Violent Recidivism

Risk Factor	Points Assigned
Age	18 to 24 = 3 points; 25 to 34 = 2 points; 35 to 44 = 1 point; Older = 0 points
Violent Appearances	0 = 0 points; 1 = 1 point; 2-3 = 2 points; 4+ = 3 points
Burglary	None = 0 points; Any = 2 points

Enter the number of points accrued above in the table below and circle the corresponding Risk Category and Label.

Points	Risk
0-1	Low
2-3	Medium
4-5	High
6 or more	Very High

RM2000/C – Risk for Sexual and Violent Recidivism Combined

Assign C-scale points from each of the V and S scale Categories.

Risk Category for S scale	Low	Medium	High	Very High
C Points assigned	0	1	2	3
Risk Category for V scale	Low	Medium	High	Very High
C Points assigned	0	1	2	3

C-scale Labels Derived From C-Scores

Circle the C-scale score and Label that applies.

Score on C-Scale	Label
0	Low
1	Medium
2	Medium
3	High
4	High
5	Very High
6	Very High



Sexual Offending: Theory, Research, and Prevention is the official journal of the International Association for the Treatment of Sexual Offenders (IATSO).



leibniz-psychology.org

PsychOpen GOLD is a publishing service by Leibniz Institute for Psychology Information (ZPID), Germany.