

Actuarial prediction of violent recidivism in mentally disordered offenders

ROBERT J. SNOWDEN^{1*}, NICOLA S. GRAY^{1,2}, JOHN TAYLOR³
AND MALCOLM J. MACCULLOCH^{1,2,4}

¹ *School of Psychology, Cardiff University, Cardiff, UK;* ² *South Wales Forensic Psychiatric Service, Caswell Clinic, Glamrhed Hospital, Mid Glamorgan, UK;* ³ *Kneesworth House Hospital, Bassingbourne-cum-kneesworth, Royston, Hertfordshire, UK;* ⁴ *Department of Psychological Medicine, University of Wales College of Medicine, Cardiff, UK*

ABSTRACT

Background. Actuarial instruments may be useful in predicting long-term violence in mentally disordered patients. We compared two instruments that differ in terms of what they are designed to predict (general *versus* violent recidivism) and the inclusion of stable mental health variables.

Method. A large sample of mentally disordered patients were scored on two risk assessment instruments, the Violence Risk Appraisal Guide (VRAG) and the Offender Group Reconviction Scale (OGRS), based on information at the point of discharge. Their criminal histories for at least 2 years following discharge were obtained from official records.

Results. Both instruments were good predictors of both violent and general offending. Over shorter periods (<1 year) the VRAG had very good predictive validities for both types of offences [areas under the receiver operating characteristic curves (AUCs) >0.85], which were significantly better than the OGRS. For longer follow-up periods the instruments had approximately equal prediction accuracy. However, both instruments predicted far more offences than were in fact recorded.

Conclusions. The VRAG is a very good predictor of future violence in the UK sample. The OGRS may also be of value as it can be completed quickly and without the need for mental health variables. Caution is needed, however, as both instruments appeared to over-predict the levels of reconvictions in this sample.

INTRODUCTION

An important task for many clinicians is to assess the probability of future violent acts in a person they are considering for discharge, and therefore how to manage this risk. The problem is difficult in that the risk assessment in the short term (days) may differ from that in the long term (years). In this paper we focus on the longer-term prediction of risk (months–years).

The unstructured clinical judgement of the future risk of an individual committing a violent act has been demonstrated to be of poor validity

in many studies (for reviews see Coccozza & Steadman, 1976; Monahan *et al.* 2001; Quinsey *et al.* 2006). Hence, instruments to aid clinical judgement were developed. In an actuarial instrument, factors known to predict a particular outcome (such as a violent act) are combined using statistically derived weightings to form a composite measure. There is strong evidence for the value of such actuarial tools in predicting future criminal behaviour (Gardner *et al.* 1996*a,b*; Grove *et al.* 2000; Monahan *et al.* 2001).

Actuarial instruments have been developed to predict general recidivism (e.g. Andrews & Bonta, 1985; Copas & Marshall, 1998) or violent recidivism (Quinsey *et al.* 1998). Those for the

* Address for correspondence: Professor R. J. Snowden, School of Psychology, Cardiff University, Cardiff CF10 3AT, Wales, UK.
(Email: snowden@cardiff.ac.uk)

prediction of general recidivism seem to have been developed for use within the correctional system for non-mentally disordered offenders (for a review see Gendreau *et al.* 1996), whereas those that aim to predict violence seem to have been developed for use in patient populations (Bonta *et al.* 1998). However, a meta-analysis of the factors that predict violent recidivism, and those that predict any recidivism, shows that the factors are very similar (Bonta *et al.* 1998; Phillips *et al.* 2005). If this is true, several conclusions follow. First, instruments developed to predict general recidivism should also be successful at predicting violent recidivism. Second, it means we can use general offending as a proxy for violent offending when testing risk assessment instruments. For instruments that are designed to predict violence it is often difficult to gather enough data on violent reconvictions because of the relative rarity of these events (particularly if using subgroups of specific offender types, etc.) and therefore long time periods and large sample sizes are required. Although such direct tests are the best form of evidence for an instrument designed to predict violence, we suggest that the ability to predict any reconviction is a useful proxy under conditions where large samples or long times are not practicable.

The role of mental health variables in predicting crime and recidivism has long been debated (e.g. Monahan, 1992; Arboleda-Florez, 1998; Appelbaum *et al.* 2000; Arseneault *et al.* 2000; Hodgins & Janson, 2002). The weight of evidence suggests that mental health variables do not play a large role in predicting long-term violent recidivism in the community for people with a history of violence (Bonta *et al.* 1998; Gray *et al.* 2004). Mental health variables do, however, play an important role in the short-term prediction of violence, such as in institutional settings (Gray *et al.* 2003; McNeil *et al.* 2003), and may well have a role in predicting the onset of violent behaviour (e.g. Fava, 1990) and/or the context in which it is observed. We should also note that many studies that failed to find effects of mental health variables examined patients discharged from hospital. In such cases the mental health problems would have been treated and this lack of an effect of mental health variables on future violence may be indicative of the successful

treatment and management of mental health issues.

To examine these issues we chose two well-established risk assessment instruments, one for violence reconvictions that incorporates mental health variables, and one for general reconviction that does not.

The Violence Risk Appraisal Guide (VRAG) (Harris *et al.* 1993; Quinsey *et al.* 1998) is one of the most well-established actuarial instruments for risk of future violence. There are several reports of its ability to predict violent recidivism in those released from a forensic setting in North America (Harris *et al.* 1993, 2002, 2003; Rice & Harris, 1995, 1997; Barbaree *et al.* 2001; Glover *et al.* 2002; Douglas *et al.* 2005); however, the agreement is not complete (Kroner & Mills, 2001; Loza *et al.* 2002). The VRAG has also shown good predictive powers in samples from Scandinavia (Grann *et al.* 2000; Tengstrom, 2001; Grann & Wedin, 2002) and in continental Europe (Pham *et al.* 2005; Urabaniok *et al.* 2006). The use of the VRAG has been extended to show that it can predict future violence in a non-forensic sample (Harris *et al.* 2004). In the UK it has been shown to have some success in predicting in-patient violence (Doyle *et al.* 2002), and a moderate effect in predicting community violence (Doyle & Dolan, 2006).

The Offender Group Reconviction Scale (OGRS) was developed by the Home Office (UK Government) for use within the probation service of England and Wales and was therefore designed for use with general offenders rather than those with a mental disorder. It was chosen for comparison to the VRAG as it was specifically designed to include only a small number (six) of easily available and unambiguously defined items that do not require any clinical opinion or judgement. The OGRS is therefore easy to use, fast, reliable, and can be used without specific professional qualifications or experience. The OGRS yields a number, indicating the probability of any reconviction within 2 years. The OGRS underwent a revision (Taylor, 1999) based on a sample of over 30 000 offenders. It was also modified so that it can make specific predictions about 'serious' reconvictions (defined as sexual or violent convictions) in those who have a previous history of such an offence. Despite this impressive construction

sample, there has been remarkably little work to exploit this instrument. In a recent paper, however, Gray *et al.* (2004) showed that the OGRS was the best predictor of subsequent reconvictions in a sample of mentally disordered offenders (see also Coid *et al.* unpublished observations), supporting the notion that factors that predict reconvictions are similar in mentally disordered and non-mentally disordered populations (Bonta *et al.* 1998).

METHOD

Design

The study was a pseudo-prospective case-note analysis of patients discharged from independent sector, medium secure facilities in the UK. The dependent variables were (1) time to offence; (2) the proportion of the sample who offended *versus* did not offend during the follow-up period; and (3) type of offence (general *versus* violent). Violent offences included all offences classified as violence against the person by the Home Office and kidnap, criminal damage endangering life, robbery, rape and indecent assault. General offences included all offences including those also classified as violent.

Participants

Patients were discharged from four independent sector, medium secure units (Llanarth Court Hospital, Kneesworth House Hospital, Stockton Hall Hospital, and Redford Lodge), between December 1992 and 30 September 2001. The total sample consisted of 996 male patients with a mean age at discharge of 37.7 years (s.d. = 9.2, range 16.9–71.2 years). Most patients (69.2%) were of Caucasian ethnic origin, 21.6% were of Black Caribbean or Black African origin, 2.4% were of Asian origin, 1.5% were of other or mixed ethnicity and 5.2% were of unknown ethnicity. The mean length of stay within the medium secure service was 436 days (s.d. = 510, range 7–3785 days).

Primary diagnosis was divided into affective disorder (9.9%), personality disorder (9.0%), schizophrenia or psychotic disorder (56.2%), drug-induced psychosis (4.7%), mental retardation (8.5%) and 'other' diagnoses (8.4%; anxiety disorder, developmental disorder, organic disorder and epilepsy), with 3.2% patients of unknown diagnosis. Diagnoses were made by

a consultant psychiatrist upon admission to hospital using ICD-10 criteria (WHO, 1992).

For various reasons it was not possible to gather exactly the same data on all the participants. Thus many of the analyses presented here are on subsamples of this overall population. For each subsample used (e.g. those with a valid VRAG score and a follow-up of 5 years or more), we compared the above patient characteristics (age, diagnosis, etc.) to the overall population. No significant differences were found.

Measures

The VRAG (Quinsey *et al.* 1998) comprises 12 items, including such items as the Psychopathy Checklist Revised (PCL-R; Hare, 2004) score (which in turn has 20 items), elementary school adjustment, offender's age at time of index offence, etc. If we could not score a particular item then that item was rated as a '0'. We note that the updated manual (Quinsey *et al.* 2006) prorating is now recommended. We did not score the VRAG if more than four items could not be scored.

Scoring of this instrument did not commence at the start of the project and hence we obtained VRAG scores from 421 patients in our sample. Comparison of those with and without VRAG scores did not reveal any significant differences in age, gender, diagnosis or outcome variables (number of reoffences, number of violent reoffences).

The OGRS (Copas & Marshall, 1998) and the revised version OGRS-2 (Taylor, 1999) are criminogenic risk assessment tools based solely upon an offender's history of offending and certain demographic variables (e.g. age and gender). The OGRS was developed for use in probation settings and does not contain any variables related to mental health. It estimates the probability that an offender will be reconvicted within 2 years of release based on six variables concerning the offender (e.g. age, gender, type of offence). The OGRS-2 uses 10 variables to calculate a score. Unless otherwise stated we do not differentiate these two versions of the instrument in this paper, and will simply refer to them as the OGRS. The OGRS score cannot be calculated in people who do not have previous convictions ($n = 198$ in our sample) and where any item cannot be scored ($n = 157$).

Table 1. *Percentage of patients reconvicted for general offences as a function of time since discharge*

OGRS score	6 months	1 year	2 years	3 years	5 years
0–20	0.7 (1/148)	2.7 (4/148)	5.4 (8/148)	8.8 (12/136)	14.3 (14/98)
21–40	3.9 (6/156)	8.3 (13/156)	14.7 (23/156)	21.1 (32/152)	33.6 (42/125)
41–60	8.3 (10/120)	14.2 (17/120)	23.3 (28/120)	31.0 (35/113)	40.7 (37/91)
61–80	7.8 (9/115)	21.8 (25/115)	40.9 (47/115)	48.1 (52/108)	62.0 (57/92)
81–100	20.0 (19/95)	41.1 (39/95)	63.8 (60/94)	74.2 (69/93)	79.8 (71/89)

Data are split according to Offender Group Reconviction Scale (OGRS) categories. Numbers in parentheses are the number of people reconvicted and the total sample size in each category.

Overall, we obtained OGRS probability scores from 641 patients in our sample, collapsed across OGRS ($n=185$) and OGRS-2 ($n=456$). Comparison of those for whom we could obtain an OGRS score against those for whom crucial data were missing did not reveal any significant differences in age, gender, diagnosis or outcome.

Procedure

Ethical approval was obtained from the Ethical Committee of the School of Psychology, Cardiff University. Four psychologists completed all the assessments by access to file-based information. Each assessor was trained on the two risk assessment instruments and did not start data collection until their inter-rater reliability (based on a 'test bank' of 20 cases) was above 0.90. On a test sample of 20 cases the raters had a collective intraclass correlation coefficient (ICC) of 0.95.

All background psychiatric and mental health reports on the patients were obtained, as were full criminal record history, admission and discharge reports, social work and probation information, and nursing records. All lifetime convictions were obtained from the Home Office Offenders' Index (a UK Government database of all convictions), both prior to admission to hospital and following discharge. Convictions following date of discharge were classified as 'outcome' data in terms of offending following release from hospital. All risk assessments were completed blind to outcome following discharge. The data available to us were the dates of any reconvictions following discharge. This is not necessarily the same as 'time at risk' because if a person is incarcerated for a minor offence this might well reduce their chances of committing another more serious crime. To

account for this problem we used a conservative measure and removed a person from further analysis at the time of their first reconviction.

RESULTS

OGRS

The 641 patients with OGRS scores were followed up for a minimum of 2 years (mean = 6.32 years, $s.d. = 2.06$, range 2.02–11.96 years). The mean number of offences in the follow-up period was 3.36 ($s.d. = 8.17$, range 0–62) and 37.9% of participants were convicted of an offence during the follow-up period. The average OGRS score was 46 ($s.d. = 27$, range 2–99).

Table 1 shows that the OGRS has good predictive abilities; for example, at the 1-year follow-up a person in the highest category is about 20 times more likely to be reconvicted than a person in the lowest category. It should be noted that the absolute levels of reconviction are somewhat lower than predicted by the OGRS score. For people in the 41–60 category, approximately 50% of them should have been reconvicted after 2 years. In fact, only around half of this expected number were reconvicted. Inspection of Table 2 shows that the OGRS also has good predictive abilities for violent crime. For example, at a 2-year follow-up people in the highest category were more than 20 times more likely to be reconvicted for a violent offence than someone in the lowest category.

The OGRS-2 (Taylor, 1999) specifically produces a separate category score for the prediction of violence and/or sexual crime. For those without a previous history of such an offence, no category is provided. Reconviction rates at 2 and 5 years are given in Table 3. The predicted probability of such convictions after 2 years as reported in the construction sample (Taylor,

Table 2. *Percentage of patients reconvicted for a violent offence as a function of time since discharge*

OGRS score	6 months	1 year	2 years	3 years	5 years
0-20	0 (0/147)	0 (0/144)	0.7 (1/141)	2.3 (3/127)	4.6 (4/88)
21-40	1.3 (2/152)	2.7 (4/147)	4.4 (6/138)	6.2 (8/129)	8.8 (8/91)
41-60	1.8 (2/112)	2.8 (3/106)	4.2 (4/95)	7.1 (6/84)	10.0 (6/60)
61-80	2.7 (3/112)	7.8 (8/103)	17.0 (16/94)	20.5 (17/83)	29.7 (19/64)
81-100	5.6 (5/89)	10.8 (9/83)	19.5 (14/77)	24.6 (17/69)	28.3 (17/60)

Data are split according to Offender Group Reconviction Scale (OGRS) categories.

Numbers in parentheses are the number of people reconvicted and the total sample size in each category.

Table 3. *Percentage of patients reconvicted for violent offences as a function of time since discharge*

OGRS-2 category	Expected ^a	Violent (2 years)	Violent (5 years)
Some risk (1-10)	6.0	1.9 (2/105)	7.6 (5/66)
Moderate (11-17)	15.0	2.6 (2/78)	6.5 (3/46)
Raised (18-24)	21.0	11.8 (6/51)	23.3 (7/30)
High (25+)	34.0	17.4 (15/86)	32.7 (18/55)

Data are split according to Offender Group Reconviction Scale (OGRS)-2 serious reconviction categories.

Numbers in parentheses are the number of people reconvicted and the total sample size in each category.

^a Based upon the predictions from data from Taylor (1999).

1999) are provided for comparison. Once again, it seems that the OGRS prediction for violent offences does very well in defining who are the most likely to reoffend violently, but overestimates this probability for mentally disordered offenders (by a factor of around 2).

VRAG

The 421 patients for whom we had VRAG scores were followed up for a minimum of 2 years (mean = 6.22 years, s.d. = 2.22, range 2.02-11.39 years). The mean number of offences in the follow-up period was 2.59 (s.d. = 6.80, range 0-50) and 32.5% of participants were convicted of an offence (general) during the follow-up period. The average VRAG score was 3.29 (s.d. = 10.35, range -24 to 36).

Tables 4 and 5 show that the VRAG has good predictive abilities for predicting both general and violent reconvictions. We compared the predicted 7-year violent reconviction rates with the longest time period we had available, namely 5 years. For the higher (and hence more

important) categories our reconviction rates are around 50% of the expected rate (discarding those cells that have very small numbers). Adjusting for the differing follow-up periods does not produce a large diminution in this discrepancy. Thus, the VRAG score was a good indicator of who is most likely to be violent, but overestimates the absolute probability of violence in this mentally disordered UK sample.

Efficacy of OGRS and VRAG to predict general and violent recidivism

The use of signal detection theory has been suggested to determine the efficacy of any risk assessment tool (Mossman, 1994; Rice & Harris, 2005). The major attraction of signal detection theory is that it is relatively immune to changes in the base rate of the behaviour to be predicted, and so the efficacy of the instrument(s) can be compared over different samples, time periods, etc. It also does not require any 'cut-off' scores to be assigned to the risk prediction instrument as all values of the instrument can be used. However, its drawback is that it requires a binary decision outcome (the person was reconvicted or not, rather than number or seriousness of reconvictions). Figure 1 shows the receiver operating characteristic (ROC) curve for the OGRS for reconviction for any (general) offence after a 2-year follow-up (solid line), while the broken line is for reconviction for a violent offence over the same period. The OGRS produced an area under the curve (AUC) of 0.785 (s.e. = 0.020) and 0.762 (s.e. = 0.035) for general and violent reconvictions respectively. The VRAG (Fig. 1) produced AUCs of 0.743 (s.e. = 0.028) and 0.776 (s.e. = 0.045) for general and violent reconvictions respectively. All of these AUCs are significantly above chance levels ($p < 0.0001$) and

Table 4. Percentage of patients reconvicted for any offence as a function of time since discharge

VRAG category	6 months	1 year	2 years	3 years	5 years
1	0 (0/2)	0 (0/2)	0 (0/2)	0 (0/2)	0 (0/2)
2	0 (0/17)	0 (0/17)	0 (0/17)	0 (0/16)	8.3 (1/12)
3	0 (0/47)	0 (0/47)	6.4 (3/47)	11.1 (5/45)	11.1 (4/36)
4	1.1 (1/91)	4.4 (4/91)	9.9 (9/91)	12.8 (11/86)	16.5 (15/67)
5	1.1 (1/93)	7.5 (7/93)	20.4 (19/93)	25.3 (22/87)	35.7 (25/70)
6	7.1 (7/98)	18.4 (18/98)	27.6 (27/98)	37.9 (33/87)	55.7 (39/70)
7	20.4 (11/54)	27.8 (15/54)	44.4 (24/54)	53.9 (28/52)	68.2 (30/44)
8	21.4 (3/14)	42.9 (6/14)	57.1 (8/14)	64.3 (9/14)	75.0 (9/12)
9	20.0 (1/5)	20.0 (1/5)	60.0 (3/5)	60.0 (3/5)	60.0 (3/5)

Data are split according to Violence Risk Appraisal Guide (VRAG) score categories (termed 'risk bins' – the higher the category, the higher the risk).
 Numbers in parentheses are the number of people reconvicted and the total sample size in each category.

Table 5. Percentage of patients reconvicted for a violent offence as a function of time since discharge

VRAG category	6 months	1 year	2 years	3 years	5 years	Predicted 7-year rate ^a
1	0 (0/2)	0 (0/2)	0 (0/2)	0 (0/2)	0 (0/2)	0
2	0 (0/17)	0 (0/17)	0 (0/17)	0 (0/16)	0 (0/12)	8.0
3	0 (0/47)	0 (0/47)	0 (0/44)	4.8 (2/42)	5.9 (2/34)	12.0
4	1.1 (1/91)	1.1 (1/88)	3.6 (3/84)	3.9 (3/78)	5.5 (3/55)	17.0
5	0 (0/92)	1.2 (1/87)	2.6 (2/77)	4.4 (3/68)	6.3 (3/48)	35.0
6	0 (0/91)	3.6 (3/83)	11.3 (9/80)	15.9 (10/63)	22.5 (9/40)	44.0
7	10.2 (5/49)	10.6 (5/47)	18.2 (8/44)	20.5 (8/39)	28.6 (8/28)	55.0
8	7.1 (1/14)	23.1 (3/13)	25.0 (3/12)	36.4 (4/11)	44.4 (4/9)	76.0
9	20.0 (1/5)	20.0 (1/5)	25.0 (1/4)	33.3 (1/3)	33.3 (1/3)	100

Data are split according to Violence Risk Appraisal Guide (VRAG) score categories (termed 'risk bins' – the higher the category, the higher the risk).
 Numbers in parentheses are the number of people reconvicted and the total sample size in each category.
^a Figures derived from Quinsey et al. (1998).

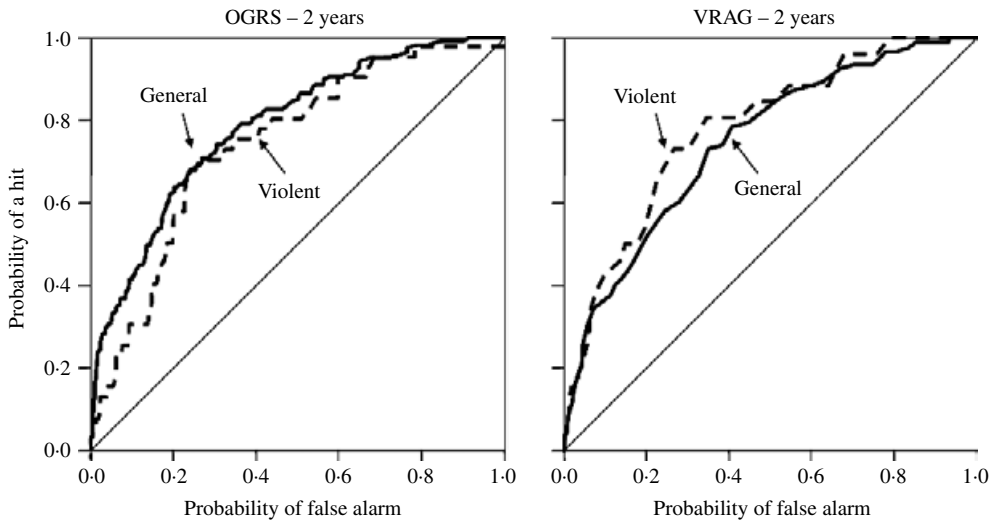


FIG. 1. The receiver operating characteristics (ROC) for the OGRS and VRAG risk assessment techniques for predicting general offending (solid line) or violent offending (broken line).

Table 6. Base rates and AUCs for the risk instruments as a function of follow-up period for general and violent reconvictions

Follow-up period	VRAG			OGRS		
	Base rate	AUC	s.e.	Base rate	AUC	s.e.
General reconvictions						
6 months	5.1	0.837 ^a	0.035	7.0	0.744	0.035
1 year	11.9	0.788	0.030	15.4	0.763	0.025
2 years	22.1	0.743	0.028	26.2	0.785	0.020
3 years	28.2	0.750	0.027	32.9	0.780	0.020
5 years	39.3	0.746	0.028	44.8	0.765	0.021
Violent reconvictions						
6 months	2.0	0.862 ^a	0.071	1.9	0.756	0.070
1 year	3.3	0.862 ^a	0.050	3.7	0.767	0.049
2 years	7.1	0.776	0.045	7.2	0.762	0.035
3 years	9.6	0.749	0.047	10.3	0.743	0.037
5 years	13.0	0.756	0.048	14.7	0.720	0.036

VRAG, Violence Risk Appraisal Guide; OGRS, Offender Group Recidivism Scale; AUC, area under the curve; s.e., standard error.

All AUCs are significant ($p < 0.001$).

^a AUCs between VRAG and OGRS differ ($p < 0.05$).

are regarded as a 'large' effect size (Rice & Harris, 2005).

Relative efficacy of each instrument to predict violent versus general reoffending

To test whether the instrument(s) were any better at predicting general or violent reoffending, ROCs were constructed and the AUCs were calculated for time periods from 6 months to 5 years (see Table 6). AUCs were compared using the methods recommended by Hanley & McNeil (1983). Note that this method requires that both instruments are scored on the same individual. We obtained 320 such cases where the VRAG and OGRS-2 were completed on the same individual, so the actual AUCs used in this statistical analysis are slightly different from the AUCs in Table 6. However, these differences in AUCs did not amount to more than 0.01 in any case.

For the VRAG, the AUCs for predicting violence are slightly larger than for general offending, particularly for short time-prediction intervals, but did not differ statistically at any follow-up period ($p > 0.05$). For the OGRS there seem to be slightly higher AUCs for general offending at the longer follow-up periods; however, these differences were not statistically significant ($p > 0.05$). Thus, both the VRAG and OGRS have similar predictive properties for general and violent reconvictions.

The OGRS-2 (Taylor, 1999) specifically produces a separate category score for the prediction of violence and/or sexual crime. We found that the OGRS category score was predictive of violent recidivism after a follow-up period of 2 years (AUC = 0.784, s.e. = 0.045, $p < 0.0001$), but this AUC did not differ significantly ($p > 0.05$) from the ability of the standard OGRS score to predict this violent recidivism (AUC = 0.762).

Comparison of VRAG versus OGRS in predicting general reoffending

Next, we compared the efficacy of the VRAG to OGRS to predict general reconvictions (Table 6). At short time periods the VRAG was a better predictor, while at longer follow-up periods (> 2 years) the OGRS seems to be a little better. Statistical comparison of AUCs at each time period showed that the VRAG was a significantly better predictor of recidivism ($p < 0.05$) for a follow-up period of 6 months, but no other differences reached significance ($p > 0.05$).

Comparison of VRAG versus OGRS in predicting violent reoffending

Finally, we made perhaps the most important comparisons, that of comparing the efficacy of prediction for violent reconvictions between the two instruments. The VRAG had a high predictive accuracy for short periods, reaching AUC = 0.86 for predicting violence over the first year, before falling to a level of around AUC = 0.75 for longer periods (a figure very similar to the construction sample for long follow-up periods; Quinsey *et al.* 1998). The AUCs for the OGRS were very stable across time. Statistical analysis revealed that the VRAG was a significantly better predictor of violent recidivism ($p < 0.05$) for follow-up periods of 6 months and 1 year, but no other differences reached significance.

DISCUSSION

We hypothesized that the factors that predict violent recidivism are very similar to those for general recidivism, and therefore that actuarial tools designed to predict one should also predict the other. We found that the instruments were able to predict violent reconvictions with

approximately the same accuracy as they could predict general reconvictions.

We did not find evidence of any differences in predictive efficacy between the instruments at follow-up periods of 2, 3 or 5 years for either violent or general reconvictions, supporting this hypothesis. We acknowledge that this evidence is weak as it is based upon a null result, and we reiterate the point made in the Introduction that this could also be indicative of the successful recognition and management of mental health problems. The VRAG was a better predictor than the OGRS of violent reconviction at periods of 1 year or less. Two points should be noted, however. First, we performed several comparisons and did not correct for this. Hence, the chance that this could be a Type 1 error is raised. Replication of this result in an independent sample is desirable. Second, the mental health variables as used in the VRAG are of a historic, or static, nature (Mills *et al.* 2003). Thus, these results should not be used to say that current mental state is not of interest in risk assessment.

Base rates and predictions

The probability of reconviction will vary according to the base rate of offending and the specific sample being measured (Douglas *et al.* 2005). In our mentally disordered UK sample, the VRAG and OGRS overpredicted the levels of reconvictions. We note that our base rate (6% reconvictions for violent offences after 2 years) is similar to other UK samples (Maden *et al.* 2004), but much smaller than the 15% in the sample of Harris *et al.* (2002). The reasons for such differences in base rates are complex and beyond the main aim of this paper. We note that difference in the definitions of the target behaviour, ability to follow-up patients to observe such behaviours, and different treatments due to mental health status and/or probation status will all influence the base rate of the target behaviour (e.g. violent reconvictions). It should also be noted that the base rates used in defining the VRAG probability bins included charges (as well as convictions) and readmission to a maximum security hospital for a violent offence. As these data were not available in our sample, it is clear that this would contribute to the lower estimate. However, it can be stated that the probability estimates provided by the VRAG

bins (or the OGRS score) are dependent upon the sample and context, and should not be used without first giving thought as to their applicability to the person and situation (see also Mills *et al.* 2005).

VRAG

The issue of whether a particular instrument works for a given population is of great practical importance. Here we show that the VRAG performs well in predicting both general, and in particular violent, recidivism in a UK sample. The efficacy of prediction seems to be somewhat greater than that reported in a recent UK sample that measured violence by self-report and collateral reports as well as official records (Doyle & Dolan, 2006), even considering similar follow-up periods (6 months). The reasons for this discrepancy are not clear but may reflect differences in demographics such as gender.

The efficacy of prediction of the VRAG for long follow-ups (>2 years) is very similar to that of previous samples (Quinsey *et al.* 1998; Grann *et al.* 2000; Kroner & Mills, 2001; Harris *et al.* 2002; Sjostedt & Langstrom, 2002), but the accuracy of the VRAG at shorter periods (within 1 year) is better than in previous studies. It is interesting to note that it is at these shorter follow-up periods that forensic mental health services and probation agencies have greater powers of social supervision, and that therefore the VRAG could be very useful in targeting limited resources to the highest risk offenders.

OGRS

We have recently shown that the ability of the OGRS to predict crime transfers well to the mentally disordered UK population (Gray *et al.* 2004). The present results confirm and extend this finding to the use of the OGRS-2 and, more importantly, indicate that the OGRS (both the original version and version 2) is capable of predicting violent crime as effectively as it predicts general reconvictions. We also note that the OGRS-2 has a specific scale for predicting 'serious crime' and we confirmed that this scale is indeed a good predictor of violent and sexual crimes in our mentally disordered offender sample. However, it did not provide a significant improvement over the standard OGRS score when used for predicting violent offending and is therefore somewhat redundant.

Limitations

Our study suffers from a number of limitations. First, our outcome variable of reconvictions is not ideal. The majority of violent acts do not result in a formal conviction, and it may be that offenders with a psychiatric history may be dealt with in a different manner to others. A study that is able to follow each patient in far greater detail is required to support the present findings (e.g. Doyle & Dolan, 2006). We note that where this has been done, the VRAG has similar prediction properties to the present findings (Harris *et al.* 2004). Second, OGRS cannot be used in patients without a criminal record, and this will limit its use in many settings in the mental health domain, where original detention may not be for criminal activities. Extension of the OGRS to encompass such individuals would be a welcome addition to the utility of this instrument. Third, even though we have examined risk prediction over several time periods, these are all fairly long. Many clinicians have a need to predict over the space of the next few hours, days and weeks, rather than months, and so other risk assessment instruments that incorporate current mental state (e.g. the 20-item, Historical, Clinical, Risk instrument HCR-20; Webster *et al.* 1997) may be of greater use for this type of assessment. For example the C-scale of the HCR-20 is found to be a good predictor of institutional violence over the next 3 months, but a poor predictor of reconvictions over a period of several years (Gray *et al.* 2003, 2004). It may still be the case that actuarial instruments are better suited for longer time periods and decisions such as whether to discharge into the community and the level of supervision and management required. Fourth, while the VRAG does use clinical information, it does so only as a historical variable; for instance, previous history of alcohol abuse, the Psychopathy Checklist-Revised (PCL-R) score, previous diagnosis of schizophrenia. This seems a sensible option for predicting violence in the longer term because current mental state may well change rapidly (but is well suited to predicting more immediate risk).

Use of actuarial risk assessment instruments

All risk assessments (whether they use actuarial methods or any other method) give a probability

of an event occurring, such as '25 people out of 100 people in this category will commit a violent act'. This often leads to the notion that the assessment does not apply to the individual but only to a group. However, it is clear that the statement above could be written as 'a person in this category has a 25% chance of committing a violent act', and hence the assessment clearly belongs to the individual. Furthermore, the same information could be used to place a person into a named category such as 'low risk' based on these statistics. Indeed, recent actuarial tools such as the Classification of Violence Risk (COVR) offer the information in all these three forms (Monahan *et al.* 2005).

The high AUCs of the VRAG and OGRS show that these two actuarial risk assessment instruments are very good at ordering people in terms of risk. However, high AUCs alone do not show that the probabilities given for each category (or risk score) are necessarily correct. In the present study the rate of violent reconviction was considerably less than that reported in the construct population of the VRAG (described earlier) and hence the probabilities produced should not be used for an individual.

Problems over whether a score can be applied to an individual, and others, have meant that the use of pure actuarial instruments for violence risk assessment has not been widespread. Instead, structured risk assessments, such as the HCR-20 (Webster *et al.* 1997), have become popular. We suggest that both structured risk assessment and actuarial risk assessments should be used in tandem. The two approaches bring different types of information to bear on the issue. Actuarial risk instruments are mechanical and inflexible in their risk assessment, but this mechanical approach gives great reliability, is free from personal bias, and overcomes the difficulty that people have in dealing with a large number of variables. They also tend to be based on static factors (although see earlier discussion), which are likely to remain relevant no matter what context the person is in at the moment. Their prognosis should not be ignored, but be used to anchor risk assessment. A high risk on such an instrument clearly raises the probability that the person will indeed be violent again in the future. However, we also need to understand 'how and when' that violence might

occur. Structured clinical assessments can then give us this risk formulation, which should include consideration of immediacy, likely victim group, key risk factors, and so on, and should certainly include mental health problems, in order to formulate risk management plans.

CONCLUSIONS

Our results confirm the utility of the VRAG and OGRS in the prediction of violent recidivism over a period of 6 months to 5 years in a large UK-based sample of mentally disordered offenders, and therefore support their addition to the tools available to clinicians working with similar samples. These actuarial measures offer a bias-free evaluation of risk and thus provide an anchor with regard to the level of risk management required for an individual and support the clinician in his/her formulation of risk.

ACKNOWLEDGEMENTS

This work was funded by a grant from Partnerships in Care plc. We thank the UK Home Office for providing us with information concerning criminal convictions in our cohort of mentally disordered offenders. We are also grateful to Suzanne Fitzgerald, Sophie MacCulloch, Helen Phillips, Caroline Durrant and Sian Chapple for data collection.

DECLARATION OF INTEREST

None.

REFERENCES

- Andrews, D. A. & Bonta, J. (1985). *LSI-R: The Level of Supervision Inventory – Revised*. Multi-Health Systems: Toronto.
- Appelbaum, P. S., Robbins, P. C. & Monahan, J. (2000). Violence and delusions: data from the MacArthur violence risk assessment study. *American Journal of Psychiatry* **157**, 566–572.
- Arhoda-Florez, J., Holley, H. & Crisanti, A. (1998). Understanding casual paths between mental illness and violence. *Social Psychiatry and Psychiatric Epidemiology* **33**, s38–s46.
- Arseneault, L., Moffitt, T. E., Caspi, A., Taylor, P. J. & Silva, P. A. (2000). Mental disorders and violence in a total birth cohort. *Archives of General Psychiatry* **57**, 979–986.
- Barbaree, H. E., Seto, M. C., Langton, C. M. & Peacock, E. J. (2001). Evaluating the predictive accuracy of six risk assessment instruments for adult sex offenders. *Criminal Justice and Behavior* **28**, 490–521.
- Bonta, J., Law, M. & Hanson, K. (1998). The prediction of criminal and violent recidivism among mentally disordered offenders: a meta-analysis. *Psychological Bulletin* **123**, 123–142.
- Cocozza, J. J. & Steadman, H. J. (1976). The failure of psychiatric predictions of dangerousness: clear and convincing evidence. *Rutgers Law Review* **29**, 1084–1101.
- Copas, J. & Marshall, P. (1998). The offender group reconviction scale: a statistical reconviction score for use by probation officers. *Applied Statistics* **47**, 159–171.
- Douglas, K. S., Yoemans, M. & Boer, D. P. (2005). Comparative validity analysis of multiple measures of violence risk in a sample of criminal offenders. *Criminal Justice and Behavior* **32**, 479–510.
- Doyle, M. & Dolan, M. (2006). Predicting community violence from patients discharged from mental health services. *British Journal of Psychiatry* **189**, 520–526.
- Doyle, M., Dolan, M. & McGovern, J. (2002). The validity of North American risk assessment tools in predicting in-patient violent behaviour in England. *Legal and Criminological Psychology* **7**, 141–154.
- Fava, M. (1990). ‘Anger attacks’: possible variants of panic and major depressive disorders. *American Journal of Psychiatry* **147**, 867–890.
- Gardner, W., Lidz, C. W., Mulvey, E. P. & Shaw, E. C. (1996a). Clinical versus actuarial predictions of violence in patients with mental illness. *Journal of Consulting and Clinical Psychology* **64**, 602–609.
- Gardner, W., Lidz, C. W., Mulvey, E. P. & Shaw, E. C. (1996b). A comparison of actuarial methods for identifying repetitively violent patients with mental illnesses. *Law and Human Behaviour* **20**, 35–48.
- Gendreau, P., Little, T. & Goggin, C. (1996). A meta-analysis of the predictors of adult offender recidivism. *What works! Criminology* **34**, 575–607.
- Glover, A. J. J., Nicholson, D. E., Hemmati, T., Bernfeld, G. A. & Quinsey, V. L. (2002). A comparison of predictors of general and violent recidivism among high risk federal offenders. *Criminal Justice and Behavior* **29**, 235–249.
- Grann, M., Belfrage, H. & Tengstrom, A. (2000). Actuarial assessment of risk for violence. Predictive validity of the VRAG and the historical part of the HCR-20. *Criminal Justice and Behavior* **27**, 97–114.
- Grann, M. & Wedin, I. (2002). Risk factors for recidivism among spousal assault and spousal homicide offenders. *Psychology, Crime and Law* **8**, 5–23.
- Gray, N. S., Hill, C., McGleish, A., Timmons, D., MacCulloch, M. J. & Snowden, R. J. (2003). Prediction of violence and self-harm in mentally disordered offenders: a prospective study of the efficacy of HCR-20, PCL-R and psychiatric symptomatology. *Journal of Consulting and Clinical Psychology* **71**, 443–451.
- Gray, N. S., Snowden, R. J., MacCulloch, S., Phillips, H., Taylor, J. & MacCulloch, M. J. (2004). Relative efficacy of criminological, clinical and personality measures of future risk of offending in mentally disordered offenders: a comparative study of HCR-20, PCL:SV and OGRS. *Journal of Consulting and Clinical Psychology* **72**, 523–530.
- Grove, W. M., Zald, D. H., Lebow, B. S., Snitz, B. E. & Nelson, C. (2000). Clinical versus mechanical prediction: a meta-analysis. *Psychological Assessment* **12**, 19–30.
- Hanley, J. A. & McNeil, B. J. (1983). A method of comparing the areas under the receiving operating characteristic curves derived from the same cases. *Radiology* **148**, 839–843.
- Hare, R. D. (2004). *The Hare Psychopathy Checklist – Revised (PCL-R)* (2nd edn). Multi-Health Systems: Toronto.
- Harris, G. T., Rice, M. E. & Camilleri, J. A. (2004). Applying a forensic actuarial assessment (the Violence Risk Appraisal Guide) to nonforensic patients. *Journal of Interpersonal Violence* **19**, 1063–1074.
- Harris, G. T., Rice, M. E. & Cormier, C. A. (2002). Prospective replication of the Violence Risk Appraisal Guide in predicting violent recidivism among forensic patients. *Law and Human Behavior* **26**, 377–394.
- Harris, G. T., Rice, M. E. & Quinsey, V. L. (1993). Violent recidivism of mentally disordered offenders: the development of

- a statistical prediction instrument. *Criminal Justice and Behavior* **20**, 315–335.
- Harris, G. T., Rice, M. E., Quinsey, V. L., Lalumière, M. L., Boer, D. & Lang, C.** (2003). A multisite comparison of actuarial risk instruments for sex offenders. *Psychological Assessment* **15**, 413–425.
- Hodgins, S. & Janson, C.-G.** (2002). *Criminality and Violence Among the Mentally Disordered: The Stockholm Metropolitan Project*. Cambridge University Press: Cambridge.
- Kroner, D. G. & Mills, J. F.** (2001). The accuracy of five risk appraisal instruments in predicting institutional misconduct and new convictions. *Criminal Justice and Behavior* **28**, 471–489.
- Loza, W., Villeneuve, D. B. & Loza-Fanous, A.** (2002). Predictive validity of the Violence Risk Appraisal Guide: a tool for assessing violent offender's recidivism. *International Journal of Law and Psychiatry* **25**, 85–92.
- Maden, A., Scott, F., Burnett, R., Lewis, G. H. & Skapinakis, P.** (2004). Offending in psychiatric patients after discharge from medium secure units: prospective national cohort study. *British Medical Journal* **328**, 1534.
- McNiel, D. E., Gregory, A. L., Lam, J. N., Binder, R. L. & Sullivan, G. R.** (2003). Utility of decision support tools for assessing acute risk of violence. *Journal of Consulting and Clinical Psychology* **71**, 945–953.
- Mills, J. F., Jones, M. N. & Kroner, D. G.** (2005). An examination of the generalizability of the LSI-R and VRAG probability bins. *Criminal Justice and Behavior* **32**, 565–585.
- Mills, J. F., Kroner, D. G. & Hemmati, T.** (2003). Predicting violent behaviour through a static-stable variable lens. *Journal of Interpersonal Violence* **18**, 891–904.
- Monahan, J.** (1992). Mental disorder and violent behavior. *American Psychologist* **47**, 511–521.
- Monahan, J., Steadman, H. J., Silver, E., Appelbaum, P. S., Robbins, P. C., Mulvey, E. P., Roth, L. H., Grisso, T. & Banks, S.** (2001). *Rethinking Risk Assessment: The MacArthur Study of Mental Disorder and Violence*. Oxford University Press: USA.
- Monahan, J., Steadman, H. J., Applebaum, P., Grisso, T., Mulvey, E. P., Roth, L. H., Robbins, P. C., Banks, S. & Silver, E.** (2005). *Classification of Violence Risk (COVR)*. Psychological Assessment Resources, Inc: Lutz, FL.
- Mossman, D.** (1994). Assessing predictions of violence: being accurate about accuracy. *Journal of Consulting and Clinical Psychology* **62**, 783–792.
- Pham, T. H., Ducro, C., Marghem, B. & Reveillere, C.** (2005). Prediction of recidivism among prison inmates and forensic patients in Belgium. *Annales Medico-Psychologiques* **163**, 842–845.
- Phillips, H., Gray, N. S., MacCulloch, S., Taylor, J., Moore, S. C., Huckle, P. & MacCulloch, M. J.** (2005). Risk assessment in offenders with mental disorders. Relative efficacy of personal demographics, criminal history, and clinical variables. *Journal of Interpersonal Violence* **20**, 833–847.
- Quinsey, V. L., Harris, G. T., Rice, M. E. & Cormier, C. A.** (1998). *Violent Offenders: Appraising and Managing Risk*. American Psychological Association: Washington, DC.
- Quinsey, V. L., Harris, G. T., Rice, M. E. & Cormier, C. A.** (2006). *Violent Offenders: Appraising and Managing Risk* (2nd edn). American Psychological Association: Washington, DC.
- Rice, M. E. & Harris, G. T.** (1995). Violent recidivism: assessing predictive validity. *Journal of Consulting and Clinical Psychology* **63**, 737–748.
- Rice, M. E. & Harris, G. T.** (1997). Cross-validation and extension of the Violence Risk Appraisal Guide for child molesters and rapists. *Law and Human Behavior* **21**, 231–241.
- Rice, M. E. & Harris, G. T.** (2005). Comparing effect sizes in follow-up studies: ROC area, Cohen's *d*, and *r*. *Law and Human Behavior* **29**, 615–620.
- Sjostedt, G. & Langstrom, N.** (2002). Assessment of risk for criminal recidivism among rapists: a comparison of four different measures. *Psychology, Crime and Law* **8**, 25–40.
- Taylor, R.** (1999). *Predicting Reconvictions for Sexual and Violent Offences using the Revised Offender Group Reconviction Scale*. Home Office Research Development and Statistics Directorate, Research Findings no. 104 (www.homeoffice.gov.uk/rds/rf1999.html).
- Tengstrom, A.** (2001). Long-term predictive validity of historical factors in two risk assessment instruments in a group of violent offenders with schizophrenia. *Nordic Journal of Psychiatry* **55**, 243–249.
- Urabaniok, F., Noll, T., Grunewald, S., Steinbach, J. & Endrass, J.** (2006). Prediction of violent and sexual offences: a replication study of the VRAG in Switzerland. *Journal of Forensic Psychiatry and Psychology* **17**, 23–31.
- Webster, C. D., Douglas, K. S., Eaves, D. & Hart, S. D.** (1997). *HCR-20: Assessing Risk for Violence (Version 2)*. Simon Fraser University: Vancouver, Canada.
- WHO** (1992). *The ICD-10 Classification of Mental and Behavioral Disorders*. World Health Organization: Geneva.