

Ethical and Practical Concerns Regarding the Current Status of Sex Offender Risk Assessment

Douglas P. Boer

Department of Psychology, University of Waikato

[Sexual Offender Treatment, Volume 3 (2008), Issue 1]

Abstract

Risk assessment strategies are becoming more complex and accurate, but remain problematic in terms of how such instruments are promoted and used. Currently, many people doing risk assessments of sex offenders are using actuarial (statistical) tests developed in other countries or on unique populations, with offenders very different from the test-development sample. This is clearly unethical and inappropriate and could result in needless prolonging of incarceration or the premature release of a dangerous individual and possible needless victimization of others. Other measures, such as the structured clinical guidelines may provide a short-term solution to the widespread use of unguided professional judgement and the lack of population-based actuarial tests.

Key Words: risk assessment, sexual offenders, actuarial, clinical

Introduction

The risk assessment of sexual and other violent offenders remains a controversial and important topic in forensic psychiatry and psychology. Academics routinely argue the strengths and weaknesses of clinical judgement versus actuarial tests versus structured professional judgement risk guideline assessment strategies. Professional assessors continue to do their best to minimize errors of overestimating or underestimating risk, as the first type of error results in the needless prolonging of incarceration for low risk offenders and the other type of error in needless risk for additional victimization of the public by high risk offenders. Unfortunately, it is my opinion that there is more argument in the literature as to what is the best type of risk assessment instrument than there is evidence of concern for public safety or about the ethical issues involved in the practice of risk assessment. It is my argument in this paper that if we are truly concerned about public safety and ethical practice, the academic arguments should be less of a focus and the main concern of all researchers and clinicians should be that of determining the safest, best model of practice to benefit both the public and the offenders themselves.

First, some definitions are needed to ensure we are all on the same wavelength. The most basic sort of risk assessment approach is clinical judgement which is unguided by risk assessment tests or professional risk assessment guidelines. This is not to say that all psychologists or psychiatrists who base their risk judgements on their clinical intuition and experience are always incorrect. Most clinicians, whether using risk assessment tests or not, base their assessment strategies and findings on clinical hypotheses and also follow diagnostic indicators on which to base their risk judgements. As a result, clinical judgement is generally significantly better than chance, but such judgement (regardless of the experience of the clinician) is much poorer in terms of overall accuracy than the actuarial tests and structured clinical guideline approaches. However, the research has also shown that clinical judgement, in general, is getting better over time. Recent meta-analyses (Hanson & Morton-Bourgon, 2004, 2007) showed that clinical judgement has vastly

improved over the years, presumably because clinicians are consumers of the research literature and as a result are paying attention to important variables related to risk such as sexual deviance and antisocial lifestyle.

Actuarial tests are statistical tests that use items that have been shown statistically to be related to risk. The items are initially selected through logical means, in that items that would be illogical are screened out to begin with. It makes sense, for example, to include the number of previous convictions for sexual offences or age of the offender in a statistical test, whereas it would not make sense to include hair color or number of tattoos. Items are also chosen for examination according to their presence in the research and clinical literature. All items are then evaluated for their individual predictive validity on a known group of offenders who have been released for a known period of time and selected items are then combined and weighted according to an algorithm that is chosen to maximize the overall predictive validity of the items when used together. Essentially, the result is an equation which post-dicts the recidivism data of the subject sample used in the study. This equation is composed of the best items which differentiate recidivists from non-recidivists which are then weighted according to the individual predictive power. The best overall equation (derived from predictive items that when combined result in increments in predictive validity) providing the best overall differentiation between such groups. The future usefulness of the resulting test (really a statistical equation) depends on the similarities between the test developer's sample and the offender being assessed. Differences in legal context, uniqueness of the offender being assessed, race and culture, amongst other variables all affect the usefulness of the actuarial test in question with one's current client. Actuarial tests yield a measure of the likelihood that an offender may commit another offence of the sort being assessed. Likelihood is an important aspect of risk, but clearly not the only aspect of risk that we are concerned about when doing a risk assessment.

Structured professional judgement (SPJs) risk assessment guideline measures are also referred to as guided clinical judgement or structured clinical guidelines. The first term seems most accurate in that (a) it is assumed the assessment is being done by a trained professional and (b) it is also true that the structured guidelines result in a risk assessment judgement, but in the final analysis (c) the professional doing the risk assessment is being asked to follow a set of guidelines (with some freedom to exercise clinical override) to structure his/her risk judgement. SPJs have been widely adopted by experienced clinicians around the world who are tasked with assessing reoffence risk of a variety of offenders because of the flexibility of the method and the fact that these instruments allow for a very wide analysis of risk including likelihood, but also imminence, lethality, victim specificity, and issues related to sexual disorders that may be reflected in criminal behavior (e.g., sadism, fetishism, partialism). To be sure there are many sorts of risk and the types of risk are compounded by the nature of the referral questions given to clinicians. For example, how likely is it that this offender will re-offend in a violent manner before the end of his sentence if he is granted parole ?

Current Status of Risk Assessment Research

It is probably accurate to say that the only things that risk assessment researchers agree upon are the above definitions regarding types of risk assessment methodologies. Adherents of any one sort of approach are often very steadfast in their admiration of their own sort of assessment practice or instruments. There are claims of superiority of one approach over the others, and both actuarial and SPJs against unstructured clinical approaches. There are actuarial tests that naively suggest the use of very few variables (sometimes only three or four) to describe the overall risk posed by an offender. No actuarial tests account for improvements with treatment, although we know that treated offenders reoffend at half the rate of untreated offenders. Similarly, few SPJ tests include variables that account for improvements with treatment. In addition, many of the SPJ tests often include

variables arguably unrelated or distally related to risk, or place equal emphasis on variables that have differing empirical relationships to risk, or have a great deal of overlap with existing tests. And, neither sort of risk assessment strategy account for decrements in reoffense risk due to aging.

In sum, there are problems with the risk assessment literature. Issues such as how to handle missing data or how to use dynamic (i.e., changeable) variables when accounting for changes as a result of treatment are rarely addressed by risk assessment test authors. Further, clinicians, trained as scientist-practitioners, must not be fooled by the claims of the research literature. All too often, the research literature is compiled by the test developers who must, to varying degrees, suffer from the allegiance effect—the tendency to find that one's data supports the supremacy of one's own test versus that of others—the researcher's equivalent of the therapist bias effect—the tendency to see improvement in a client where there is none. We see this effect in action when a researcher discounts data that supports a competitor's test or somehow massages the data or the presentation of meta-analytic data to show that his or her test magically outperforms competitor's tests. The result of the allegiance effect in test research is not restricted to getting one's data published when it may not be warranted, but if grandiose conclusions dissuade test users from using valid methodology, that researcher may be unintentionally cause bad clinical practice with potentially catastrophic consequences to potential victims.

In addition to above warnings, it is my contention that we must be very careful to not conclude that any test or set of guidelines can replace our clinical acumen, but to be aware that the opposite conclusion appears true as well—our clinical acumen is no match for the tests. Paradoxical as the last sentence sounds (and is on a superficial level), the essence of the dilemma is that a good risk assessment needs both our clinical wisdom and the input from both actuarial and SPJ instruments. When a clinician uses a well-known test such as the MMPI in a mental health assessment, he or she also uses their clinical observation skills and diagnostic acumen to come up with a complete clinical picture of the patient. If the clinician's observations do not match the results of the MMPI profile, the clinician has been trained to form their clinical opinion primarily based on their observations and make reference to the differing test data. Risk assessors should do the same. There is no point in blind allegiance to one's clinical intuition or hypothesis-driven diagnosis of risk if the research literature suggests that actuarial or SPJ results provide more accurate depictions of risk. However, it would seem equally foolhardy to abandon one's clinical hypotheses when dealing with clients who are each unique and may be quite different from the standardization sample of the most useful actuarial test or have idiosyncratic risk factors that are not captured in the most relevant SPJ. In fact there is research that shows that highly dangerous offenders may be underdiagnosed for risk with actuarial or SPJs.

Ethical Concerns

The area of risk assessment of any violent offender group involves many ethical issues, which are broadly, the rights of the victims and the public for safety, and the rights of the offender for accuracy and fairness. These two broad concerns are linked since a fair and accurate risk assessment ought to serve to protect the public, reduce victimization and provide an accurate picture of the offender's risk to the public so that the agencies charged with his rehabilitation or incarceration can do these tasks in the most fair and expedient fashion.

The ethical concerns for the offender and the public involve the problems of false negative (FN) and false positive (FP) errors. These concerns are, respectively, inaccurate risk depictions of the offender as not dangerous when he is (FN) and risk depictions of the offender as dangerous when he is not (FP). A FN error puts the public at risk for victimization as it may allow the offender to be released too early or not be supervised adequately upon release as the risk assessment

under-represents his risk to the public. A FP error keeps the offender in jail too long or puts too many resources into his supervision unnecessarily as the risk assessment has over-estimated his risk to the public. Neither type of risk is good for the offender or the public, but the latter type of error is preferable as our key concern is public safety and if an error is to be made it ought to be made in favor of protecting victims not offenders. However, over-incarceration is a drain on public funds and it is obviously preferable to be as accurate as possible, and release offenders when optimal so that these men (or women) can return home and become contributors to society as opposed to a burden on the taxpayer.

Perhaps the most contentious ethical issue I see emerging from the risk assessment literature is the untoward effect that over-zealous conclusions can have on the test user. Hanson and Morton-Bourgon's conclusions have changed markedly from 2004 to 2007 and some positive findings for SPJ instruments have been discounted as statistical fluke in the newer of their papers. While these authors have changed their decision rules, the effect on practice could be remarkable and unfounded. Framing opinion as fact is unethical junk science junk meta-analysis in this case. Clearly the 2007 meta-analysis doesn't clarify or help the practitioners doing the work and that should be the point of research in this area helping the clinicians who do the work of therapy and risk assessment. Until we actually have better and more consistent data, the use of a convergent approach still seems more ethically palatable than suggesting the abandonment of one assessment strategy in favor of another.

Current Recommendations the Brazilian example

In the Brazilian context it would safe to assume that an actuarial test developed elsewhere needs to be validated in Brazil before use as anything other than an indicator of potential relative risk. Actuarial tests are only completely useful when validated in the setting in which they are to be used. It makes absolutely no sense to do otherwise as such tests are derived and have any claim to predictive accuracy only when used with members of a population wherein which such tests have been evaluated. Doing otherwise is a nonsensical application, much like using influenza vaccine developed Canada to treat influenza in Brazil the strains of influenza are different and as a result the vaccine from Canada may have no effect and may even be worse than no vaccine at all. Similarly, an actuarial test may yield meaningless results, or worse, underestimate risk and result in a dangerous offender gaining the opportunity to create new victims.

The items that comprise Structured Clinical Guideline instruments are derived from the research literature and the overall test is then applied to the individual. Some of the SPJs have been shown to have validity in a variety of countries, but this does not necessarily mean that a commonly used SPJ would work in the Brazilian context. However, the items are less sample-dependent than are those in an actuarial test. Hence, I would cautiously recommend the adoption of the Sexual Violence Risk 20 (SVR-20; Boer, Hart, Kropp, & Webster, 1997), already proven to work in number of countries, as a basic SPJ for examination. However, I would also propose that the best of the actuarial measures, the STATIC-99 by Hanson and Thornton (1999) be examined for adoption. As I have written elsewhere, it may well be that a convergent appraisal of risk coming at the risk picture of a client from a number of directions may well provide the best and most well-rounded appraisal of risk. In my view, the better the risk picture, the more likely we are to come up with effective risk management strategies for that individual offender and isn't that the entire point?

I would like to repeat the five basic questions for any risk assessment of a sexual offender as recommended in the SVR-20 manual, followed by some advice in brackets. These are:

1. What is the likelihood that the individual will engage in sexual violence, if no efforts are made to manage risk? (this question is probably best answered by an appropriately validated actuarial risk instrument).
2. What is the probable nature, frequency, and severity of any future sexual violence? (this question is probably best answered by the SVR-20 which will help guide the clinician's appraisal of the offender).
3. Who are the likely victims of any future sexual violence? (this question is probably best answered by the offender's pattern of offending—an issue that is not directly asked by the SVR-20, but can be addressed via the specification of offence scenarios).
4. What steps could be taken to manage the individual's risk for sexual violence? (this question is best answered once the dynamic or changeable risk factors have been specified—presumably the best management strategies would be those that address the dynamic factors related to the offender's offending pattern—this can also be addressed via the depiction of risk-decreasing scenarios—the situations, feelings, thoughts and behaviors associated with decreasing risk potential).
5. What circumstances might exacerbate the individual's risk for sexual violence? (this question is probably best answered following completion of the SVR-20 and the depiction of risk-increasing scenarios—the situations, feelings, thoughts and behaviors associated with increasing risk potential).

Conclusions

There is much that we now know about risk assessment with sexual offenders. However, I would dispute the claims of many researchers and test developers and acknowledge that there is also a great deal that we are unsure about and as a result we need to assess risk and make our recommendations regarding case management with caution. I am not a liberal and would prefer convicted offenders to do more time in jail than harm another person. However, I also would like our risk assessment research to advance to the point where those men who are safe to release are released and reduce the cost of jailing offenders unnecessarily.

Regardless of the state of the risk assessment literature and the ethical issues involved, the Courts and Parole Boards around the world demand risk assessments on which to base their decisions. It is our responsibility to provide our best risk formulation on the individual patient while acknowledging the limitations of our field. Further, it is also our responsibility to advise policy and law makers regarding effective risk methodologies and to encourage research in this regard. Clearly we have a responsible job, as not only do we provide the basis for decisions that affect the offender's life, but also decisions that could affect the life of potential victims. As a result, we need to do the best and most ethical job possible to allow the offender to regain a normal life as soon as possible while protecting the public.

References

1. Boer, D.P. (2006). Sexual offender risk assessment strategies: is there a convergence of opinion yet? [Sexual Offender Treatment, 1](#), 1-4.
2. 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. Vancouver, B.C.: The Mental Health, Law, and Policy Institute.
3. Hanson, R.K. & Morton-Bourgon, K. (2004). Predictors of sexual recidivism: An updated meta-analysis. Public Works and Government Services Canada. Cat. No.: PS3-1/2004-2E-PDF. ISBN: 0-662-36397-3.

4. Hanson, R.K. & Morton-Bourgon, K. (2007). The accuracy of recidivism risk assessments for sexual offenders : A meta-analysis (User report 2007-1). Ottawa : Department of Public Safety and Emergency Preparedness.
 5. Hanson, R.K., & Thornton, D. (1999). Static 99: Improving actuarial risk assessments for sex offenders (User report 1999-02). Ottawa: Department of the Solicitor General of Canada.
-

Author address

Douglas P. Boer

Department of Psychology

University of Waikato

Private Bag 3105, Hamilton

New Zealand

 drdoug@waikato.ac.nz