Current Directions in Violence Risk Assessment

Jennifer L. Skeem
University of California, Irvine

John Monahan
University of Virginia School of Law

March 2011

Current Directions in Psychological Science, in press.

Current Directions in Violence Risk Assessment

Jennifer L. Skeem
University of California, Irvine

John Monahan
University of Virginia

Note: We thank Paul Appelbaum, Christopher Baird, and Daniel Murrie for helpful comments on an earlier version of this article. We disclose that one of the authors (JM) has an ownership interest in the Classification of Violence Research (COVR).

Contact:

Jennifer Skeem, Ph.D.
Professor of Psychology & Social Behavior
University of California, Irvine
4322 Social & Behavioral Sciences Gateway
Irvine, California 92697-7085
Voice: (949) 824-3782
Internet: skeem@uci.edu

Abstract

A variety of instruments have been published over recent years that improve clinicians' ability to forecast the likelihood that an individual will behave violently. Increasingly, these instruments are being applied in response to laws that require specialized risk assessments. In this article, we present a framework that goes beyond the "clinical" and "actuarial" dichotomy to describe a continuum of structured approaches to risk assessment. Despite differences among them, there is little evidence that one validated instrument predicts violence better than another. We believe that these group-based instruments are useful for assessing an individual's risk, and that an instrument should be chosen based on an evaluation's purpose (i.e., risk assessment vs. risk reduction). The time is ripe to shift attention from predicting violence to understanding its causes and preventing its (re)occurrence.

Forensic psychology has become recognized as a specialty practice area and has grown tremendously over recent years as an assessment-focused enterprise. A variety of instruments have been published to help clinicians evaluate legally relevant questions about justice-involved individuals. Many of these instruments improve clinicians' ability to forecast the likelihood that an individual will behave violently. Increasingly, these instruments are being applied in response to statutes and regulations that require specialized assessments to identify "high risk" individuals for detention or "low risk" individuals for release.

In this article, we provide a current snapshot of the violence risk assessment field. After highlighting the contexts in which risk is assessed, we describe a framework for understanding alternative approaches to assessing risk and compare them. We draw attention to modern debates about whether group-based instruments are useful for assessing an individual's risk, and whether the pursuits of risk assessment and risk reduction should be separated.

We wish to be clear about our use of terminology. We endorse the general definition of risk assessment given by Kraemer et al. (1997): "The process of using risk factors to estimate the likelihood (i.e., probability) of an outcome occurring in a population." These authors define a risk factor as a correlate that precedes the outcome in time, with no implication that the risk factor and outcome are causally related. Our outcome of focus is physical violence to others.

Legal Context

The populations in which violence risk is assessed vary across many legal contexts. In the criminal and juvenile justice systems, risk assessment can be a component of decision-making regarding bail, sentencing, and parole. In the mental health system, civil commitment on the ground of "dangerousness," commitment as a sexually violent predator, and the tort liability of clinicians for their patients' violence often turn on issues of risk assessment. Risk assessment for workplace violence and violent terrorism are also becoming increasingly common.

The law regulating the process of violence risk assessment has become much more developed in the U.S. in recent years. Some cases specify risk factors that may and may not be used to estimate risk (e.g., race is Constitutionally proscribed as a risk factor, whereas gender and age are permitted (Monahan, 2006)). Some statutes have come to explicitly require that specific instruments be administered in the risk assessment process. For example, Virginia's Sexually Violent Predator statute not only mandates the use of a specific instrument, but also specifies the cut-off score on that instrument that must be achieved to proceed further in the commitment process.

Assessment Approaches

No distinction in the history of risk assessment has been more influential than Paul Meehl's (1954) cleaving the field into "clinical" and "actuarial" (or statistical) approaches. In recent years, however, a plethora of instruments has been published that are not adequately characterized by a simple clinical-actuarial dichotomy. Rather, the risk assessment process now exists on a continuum of rule-based structure, with completely unstructured ("clinical") assessment occupying one pole of the continuum, completely structured ("actuarial") assessment occupying the other pole, and several forms of partially-structured assessment lying between the two.

The violence risk assessment process, in this regard, might usefully be seen as having the four components shown in Table 1: (1) identifying empirically-valid risk factors, (2) determining a method for measuring ("scoring") these risk factors (3) establishing a procedure for combining scores on the risk factors, and (4) producing an estimate of violence risk. It is possible to array five current approaches to violence risk assessment according to whether the approach structures (i.e., specifies rules for generating) none, one, two, three, or all four components of this process. Purely "clinical" risk assessment structures *none* of the components. The clinician selects, measures, and combines risk factors, and produces an estimate of violence risk, as his or her clinical experience and judgment indicate.

Performing a violence risk assessment by reference to a standard list of risk factors that have been found to be empirically valid (e.g., age, past violence), such as the lists provided in psychiatric texts structures *one* component of the process. Such lists function as an *aide mémoire* to identify which risk factors the clinician should attend to in conducting his or her assessment, but they do not further specify a method for measuring these risk factors.

The "structured professional judgment" (SPJ) approach exemplified by the HCR [Historical-Clinical-Risk Management]-20 (Webster, Douglas, Eaves, & Hart, 1997) structures *two* components of the process: the identification and measurement of risk factors, which may be scored as 0 if absent, 1 if possibly present or 2 if definitely present. Structured professional judgment instruments do not go further and structure how the individual risk factors are to be combined in clinical practice. As Webster et al (1997, p. 22) have stated: "it makes little sense to sum the number of risk factors present in a given case...[I]t is both possible and reasonable for an assessor to conclude that an assessee is at high risk for violence based on the presence of a single risk factor."

Approaches to risk assessment that structure *three* components of the process are illustrated by the Classification of Violence Risk (COVR; Monahan et al, 2001) and the Level of Service Inventory (LSI; Andrews, Bonta, & Wormith, 2004). These instruments structure the identification, measurement, and combination of risk factors (via a classification tree design or summing scores). But those who developed the instruments do not recommend that the final risk assessment reflect *only* the combined scores on the assessed risk factors. Given the possibility that rare factors influence the likelihood of violence in a particular case—and that, precisely because such factors rarely occur, they will never appear on an actuarial instrument—a professional review of the risk estimate is advised (while realizing that clinicians may overidentify "rare" factors). However, little is known about *how often* or *how much* clinicians modify actuarial estimates, or about the *justification* they provide for such modifications.

The best known forensic instrument that structures all *four* of the components of the violence risk assessment process is the Violence Risk Appraisal Guide (VRAG; Quinsey, Harris, Rice, & Cormier, 2006). This instrument not only structures the identification, measurement, and combination of risk factors. It also specifies that once an individual's violence risk has been actuarially characterized, the risk assessment process is complete. As Quinsey et al have stated, "What we are advising is not the *addition* of actuarial methods to existing practice, but rather the *replacement* of existing practice with actuarial methods" (p. 197).

Does One Approach Predict Better than Another?

Of these five approaches, the unstructured ("clinical") one rests upon the least empirical support. In the last major study of this approach, Lidz, Mulvey, and Gardner (1993) concluded that "clinical judgment has been undervalued in previous research. Not only did the clinicians pick out a statistically more violent group, but the violence that the predicted group committed was more serious than the acts of the comparison group. Nonetheless, the low sensitivity and specificity of these judgments show that clinicians are relatively inaccurate predictors of violence" (p. 1010).

We know of no research that systematically compares the predictive utility of strategies that structure none, one, two, three, or all four components of the process. Relevant data are available, however, on approaches that structure two or more components. Recent debates about whether it is more appropriate to structure clinical judgment (e.g., HCR-20) or replace it altogether (e.g., VRAG) has prompted a number of horse races that compare the predictive efficiency of leading instruments.

Taken together, these studies provide little evidence that one validated instrument predicts violence significantly better than another. In a recent meta-analysis of 28 studies that controlled well for methodological variation, Yang, Wong, and Coid (in press) found that the predictive efficiencies of nine risk assessment instruments (including the HCR-20, LSI-R and VRAG) were essentially "interchangeable," with point estimates of accuracy falling within a narrow band (AUC=.65 to .71). Although most of these studies used total *scores* on the HCR-20 rather than structured clinical

judgments (low/medium/high-risk), there is some evidence that those judgments both predict violence and add incremental predictive utility to scores derived by actuarially combining items (see Heilbrun, Douglas, & Yasuhara, 2010). But this latter claim is contested (Hanson & Morton-Bourgon, 2009).

Why might well-validated instruments perform equally in predicting violence? One persuasive explanation is that they tap – albeit in different ways -- "common factors" or shared dimensions of risk, despite their varied items and formats. In a innovative demonstration, Kroner, Mills, and Reddon (2005) printed the items of four well-validated instruments (e.g., LSI-R, VRAG) on strips of paper, placed the strips in a coffee can, shook the can, and then randomly selected items to create four new tools. The authors found that the "coffee can instruments" predicted violent and nonviolent offenses as well as the original instruments. Factor analyses suggested that the instruments tap four overlapping dimensions: criminal history, an irresponsible lifestyle, psychopathy and criminal attitudes, and substance abuse-related problems. Despite surface variation, well-validated instruments may generally tap "a longstanding pattern of dysfunctional and aggressive interpersonal interactions and antisocial and unstable lifestyle that are common to many perpetrators of violence" (Yang et al., in press).

The strongest risk factors for violence seem to be shared not only among risk assessment instruments, but also across key groups. In particular, an increasing body of research suggests that only a small proportion of violence committed by people with major mental illness—perhaps as little as 10 percent—is directly caused by symptoms (see Skeem, Manchak, & Peterson, in press). Most people with mental illness share leading risk factors for violence with their healthy counterparts.

Are Empirically Based Instruments Useful for Individuals?

One issue that has generated controversy is the argument of Hart, Michie, and Cooke (2007) that the margins of error surrounding individual risk assessments of violence are so wide as to make such predictions "virtually meaningless" (p. 263). Cooke and Michie (2010) concluded, "it is clear

that predictions of future offending cannot be achieved, with any degree of confidence, in the individual case" (p. 259).

This position has been vigorously contested.. For example, Hanson and Howard (2010) demonstrate that the wide margin of error for individual risk assessments is a function of having only two possible outcomes (violent/not violent) and therefore conveys nothing about the predictive utility of a risk assessment tool. Because all violence risk assessment approaches, not just actuarials, yield some estimate of the likelihood that a dichotomous outcome will occur, none are immune from Hart et al.'s (2007) argument. Instead, their argument "if true,... would be a serious challenge to the applicability of any empirically based risk procedure to any individual for anything" (Hanson & Howard, 2010, p. 277).

Our view is that group data theoretically can be, and in many areas empirically are, highly informative when making decisions about individual cases. Consider two examples from other forms of risk assessment. In the insurance industry, "until an individual insured is treated as a member of a group, it is impossible to know his expected loss, because for practical purposes that concept is a statistical one based on group probabilities. Without relying on such probabilities, it would be impossible to set a price for insurance coverage at all" (Abraham, 1986, p. 79). In weather forecasting, a wealth of data is available on given events occurring under specified conditions. Therefore, when meteorologists "predict a 70 percent chance of rain, there is measurable precipitation just about 70 percent of the time" (National Research Council, 1989, p. 46). Finally, consider the revolver analogy of Grove and Meehl (1996, p. 305-306):

> Suppose you are a political opponent held in custody by a mad dictator. Two revolvers are put on the table and you are informed that one of them has five live rounds with one empty chamber, the other has five empty chambers and one live cartridge, and you are required to play Russian roulette. If you live, you will go free. Which revolver would you choose? Unless you have a death wish, you would choose the one with the five empty chambers.

Why? Because you would know that the odds are five to one that you will survive if you pick that revolver, whereas the odds are five to one you will be dead if you choose the other one. Would you seriously think, "Well, it doesn't make any difference what the odds are. Inasmuch as I'm only going to do this once, there is no aggregate involved, so I might as well pick either one of these two revolvers; it doesn't matter which"?

Although the probabilities associated with risk assessment clearly will never be as certain as those associated with bullets being in firing chambers, we find compelling Grove and Meehl's point that group data can powerfully inform individual assessments of violence risk.

Should Risk Assessment and Reduction be Separated?

In the U.S., correctional agencies that manage a staggering number of youth and adults are increasingly endorsing structured risk assessment approaches and programs that aim to reduce reoffending by targeting risk factors like anger, poor self control, and antisocial attitudes. In this context, companies have begun marketing complex (and poorly-validated) assessment systems that explicitly include purported treatment-relevant variables in their risk estimates and ostensibly serve the risk reduction enterprise better than simple actuarial tools. Theoretically, treatment-relevant variables are risk factors that can be changed and are causally linked with violence.

This has sparked debate about whether the pursuit of risk assessment and risk reduction should be separated or integrated. Baird (2009) favors separation, arguing that the addition of treatment-relevant variables to otherwise parsimonious risk equations that emphasize past (mis)behavior will dilute their predictive utility. Andrews (2009) challenges Baird's data and reasoning, arguing that some treatment-relevant variables are risk factors and should be integrated in risk estimates. He argues that efficient prediction can be achieved by statistically selecting and combining a few highly predictive risk factors, but tools that sample risk domains more broadly and include treatment-relevant risk factors can be equally predictive.

Given a pool of instruments that are well-validated for the groups to which an individual belongs, our view is that the choice among them should be driven by the ultimate purpose of the evaluation. If the ultimate purpose is to characterize an individual's likelihood of future violence relative to other people, then choose the most efficient instrument available. This is appropriate for a single event decision in which there is no real opportunity to modify the risk estimate based on future behavior (see Heilbrun, 1997). If the ultimate purpose is to manage or reduce an individual's risk, then value may be added by choosing an instrument that includes treatment-relevant risk factors. (Although an integrated instrument would be most parsimonious, we can easily envision a two-stage process in which a risk-assessment step was followed by an independent risk-management step.) This choice is appropriate for ongoing decisions where the risk estimate can be modified to reflect ebbs and flows in an individual's risk over time. Beyond focusing risk reduction efforts, these instruments could provide incentive for changing behavior (a parole board cannot advise an inmate to ''undo'' his past commission of an assault, but can advise him to develop employment skills).

This view comes with three important caveats. First, techniques that include treatment-relevant risk factors will add no value to simpler approaches unless the risk assessment is followed by a period of control over the individual during which those factors are translated into an individual supervision and treatment plan (rather than simply filed away), and systematically targeted with appropriate services (rather than ignored in resource allocation). Second, treatment-relevant variables can and do appear in statistically derived risk assessment instruments (see Monahan et al., 2001); an instrument's degree of structure cannot be equated with its relevance to risk reduction. Third, even well-validated instruments offer little direct validity data for the treatment-relevant variables they include. It is not enough to demonstrate that a variable is a risk factor for violence; here, it must further be shown that the variable reduces violence risk when successfully changed by treatment (i.e., is a *causal* risk factor (Kraemer et al, 1997)). This is a crucial issue to address in future research, if tools continue to be sold on the promise of informing risk reduction.

Future Directions

The violence risk assessment field may be reaching a point of diminishing returns in instrument development. We might speculate that incremental advances could be made by exploring novel assessment methods, including implicit measures (Nock et al., 2010), or simple heuristics (Goldstein & Gigerenzer, 2009). But specific structured techniques seem to account for very little of the variance in predictive accuracy. If we are approaching a ceiling in this domain, there clearly are miles to go on the risk reduction front. We hope that forensic psychology shifts more of its fruitful attention from predicting violence to understanding its causes and preventing its (re)occurrence.

References

Abraham, K. (1986). *Distributing Risk: Insurance, Legal Theory, and Public Policy.* New Haven: Yale University Press.

Andrews, D. (2009, Sept.). The Level of Service Assessments: A question of confusion, selectivity, and misrepresentation of evidence in Baird (2009). Paper presented at the meeting of the International Community Corrections Association (Orlando, Florida). Accessed June 30, 2009 at: http://www.iccaweb.org/documents/D.Andrews_Service.Assessments.pdf

Andrews, D., Bonta, J., & Wormith, J. (2004). *Manual for the Level of Service/Case Management Inventory (LS/CMI).* Toronto: Multi-Health Systems.

Baird, C . (2009). *A question of evidence: A critique of risk assessment models used in the justice system. Madison: National Council on Crime & Delinquency.* Accessed June 30, 2010 at www.nccd-crc.org/nccd/pubs/2009

Cooke, D., & Michie, C. (2010). Limitations of diagnostic precision and predictive utility in the individual case: A challenge for forensic practice. *Law and Human Behavior, 34,* 259-274.

Goldstein, D., & Gigerenzer, G. (2009). Fast and frugal forecasting. *International Journal of Forecasting, 25,* 760-772.

Grove, W.M. & Meehl, P.E. (1996). Comparative efficiency of informal (subjective, impressionistic) and formal (mechanical, algorithmic) prediction procedures: The clinical-statistical controversy. *Psychology, Public Policy, and Law, 2,* 293-323.

Hanson, R.K., & Howard, P. (2010). Individual confidence intervals do not inform decision-makers about the accuracy of risk assessment evaluations. *Law and Human Behavior, 34,* 275-281.

Hanson, R. K., & Morton-Bourgon, K. E. (2009). The accuracy of recidivism risk assessments for sexual offenders: A meta-analysis. *Psychological Assessment, 21,* 1-21.

Hart, S. D., Michie, C., & Cooke, D. J. (2007) Precision of actuarial risk assessment instruments: Evaluating the 'margins of error' of group versus individual predictions of violence. *British Journal of Psychiatry, Supplement, 49, Vol 190 60-65.*

Heilbrun, K., Douglas, K. S, & Yasuhara, K. (2009). Controversies in violence risk assessment. In J. L. Skeem, K. S. Douglas, & S. O. Lilienfeld (Eds), Psychological science in the courtroom: Controversies and consensus (pp. 333-356). New York, NY: Guilford.

Heilbrun, K. (1997). Prediction versus management models relevant to risk assessment: The importance of legal decision-making context. *Law & Human Behavior, 21,* 347–359.

Kraemer, H., Kazdin, A., Offord, D., Kessler, R., Jensen, P., & Kupfer, D. (1997). Coming to terms with the terms of risk. *Archives of General Psychiatry, 54,* 337-343.

Kroner, D., Mills, J., & Morgan, B. (2005). A coffee can, factor analysis, and prediction of antisocial behavior: The structure of criminal risk. *International Journal of Law & Psychiatry, 28,* 360-374.

Lidz, C., Mulvey, E., & Gardner, W. (1993). The accuracy of predictions of violence to others. *Journal of the American Medical Association, 269,* 1007-1011.

Meehl, P. (1954). *Clinical Versus Statistical Prediction: A Theoretical Analysis and a Review of the Evidence.* Minneapolis: University of Minnesota.

Monahan, J. (2006). A jurisprudence of risk assessment: Forecasting harm among prisoners, predators, and patients. *Virginia Law Review, 92*, 391-435.

Monahan, J., Steadman, H., Silver, E., Appelbaum, P., Robbins, P., Mulvey, E., Roth, L., Grisso, T., and Banks, S. (2001). *Rethinking Risk Assessment: The MacArthur Study of Mental Disorder and Violence*. New York: Oxford University Press.

National Research Council (1989). *Improving Risk Communication*. Washington DC: National Academy Press.

Knock, M.K., Park, J.M., Finn, C.T., Deliberto, T.L., Dour, H.J, and Banaji, M.R. (2010, April). Measuring the suicidal mind: Implicit cognition predicts suicidal behavior. *Psychological Science,* 511-517.

Skeem, J., Manchak, S., & Peterson, J. (in press). Correctional policy for offenders with mental disorder: Creating a new paradigm for recidivism reduction. *Law and Human Behavior.*

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

Webster C, Douglas K, Eaves D, and Hart, S. (1997). *HCR-20: Assessing Risk for Violence (Version 2).* Vancouver, Simon Fraser University.

Yang, M., & Wong, S. C. P., & Coid, J. W. (in press). The efficacy of violence prediction: A meta-analytic comparison of nine risk assessment tools. *Psychological Bulletin.*

Table 1. Increasingly Structured Violence Risk Assessment Approaches

Approach/Tool	Structured Component of the Violence Risk Assessment Process			
	Identify risk factors	Measure risk factors	Combine risk factors	Produce final risk estimate
Clinical judgment				
Standard list of risk factors	X			
HCR-20	X	X		
COVR & LSI-R	X	X	X	
VRAG	X	X	X	X