Risk Assessment Tool Literature Review

Prepared by

Megan E. Collins

Department of Criminology and Criminal Justice

University of Maryland, College Park

With the Assistance of

Professors Brian Johnson, Thomas Loughran, James Lynch and Kiminori Nakamura, Department of Criminology and Criminal Justice, University of Maryland, College Park

September 30, 2014

Overview of Actuarial Risk Assessment

Purpose of Risk Assessment

Risk assessment, which involves prospectively estimating the probability a person will reoffend, is not a new concept in criminal justice. Discussion about offender risk emerged with the advent of parole and probation systems in the nineteenth century. More recently, risk assessment has been used in bail and pretrial release assessments, probation decisions, and in predicting future behavior of parolees (Goldkamp & Gottfredson, 1985; Champion 1994; Palacios, 1994). The use of risk in structuring sentences largely decreased in popularity in the 1970s following the introduction of truth in sentencing laws and retributive sentencing. However, the current economic and political climate—largely driven by issues of fiscal constraints, limited resources, and human rights issues—has put increased pressure on states to control prison populations and to reconsider the use of erstwhile risk-based policies, such as selective incapacitation (Monahan & Skeem, 2014).

As of April 1, 2014, the Maryland Department of Public Safety and Correctional Services was housing 21,149 inmates in state facilities.[1] The financial, social, and ethical implications of this undertaking are enormous. Today the Maryland criminal justice system has the opportunity to more effectively identify which offenders can be given non-custodial sentences without

[1] "Quarterly DPSCS Inmate Characteristic Report: April 2014"
https://www.dpscs.state.md.us/publicinfo/pdfs/stats/final/inmate-char_reports14.shtml

compromising public safety. By introducing formal risk assessment into the sentencing process, the state could reduce its correctional burden by better identifying suitable candidates for diversion sentences, and using incarceration as a last resort. A wealth of research suggests that risk assessment tools can more effectively identify those at high-risk for recidivating than professional judgment alone (e.g. Latessa & Lovins, 2010; Ægisdóttir et al., 2006; Andrews et al., 2006; Grove et al., 2000). These risk assessment tools are often paired with needs assessments, which identify criminogenic factors (those that increase crime/criminality) and inform how supervision, programming, and interventions should be applied to meet those needs (Hyatt et al., 2011). By using evidence-based risk and needs assessments to identify low-risk offenders for non-incarceration sentences, the State of Maryland has the potential to simultaneously conserve resources, ensure appropriate treatments are being assigned, and potentially even reduce the problem of recycling offenders through the criminal justice system.

Risk-based decisions made at every point in the criminal justice process may be informed by a number of factors, including a professional's experience and intuition, lengthy interviews, individual assessments, or by using actuarial risk assessments.[2] The actuarial risk assessment instruments are scientifically rigorous statistical tools typically found to be both more consistent and more accurate in their predictions than human beings (Hyatt et al, 2011; Silver & Miller, 2002; Gottfredson & Moriarty, 2006). For that reason, many jurisdictions have begun to integrate actuarial models into decision making processes in recent years. These instruments can compensate for some of the limitations of making sentencing decisions based exclusively on

[2] Actuarial assessment refers to a formal methodology that provides "a probability, or expected value, of some outcome. It uses empirical research to relate numerical predictor variables to numerical outcomes. The *sine qua non* of actuarial assessment involves using an objective, mechanistic, reproducible combination of predictive factors, selected and validated through empirical research, against known outcomes that have also been quantified" (Heilbrun, 2009: 133).

clinical judgment, such as issues with quantification, bias, limited time and resources (necessary for interview based assessments), and the required honesty and cooperation of the defendant (VanNostrand & Lowenkamp, 2013). Advocates for an actuarial approach argue that switching from a wholly clinical judgment model to a more data-driven approach could increase public safety by allowing jurisdictions to effectively incapacitate the highest-risk offenders, and consequently use public resources more efficiently (VanNostrand & Lowenkamp, 2013; Casey et al., 2011). For example, when high-risk offenders are released to the community while those rated as low-risk remain incarcerated, innocent members of the public are in unnecessary danger of victimization; additionally, incarcerating high rates of low-risk offenders places excess strain on budgets, resources, and communities (VanNostrand & Lowenkamp, 2013). Using risk assessment to identify low-risk offenders instead allows community programming to be utilized more efficiently, saves prison beds for the most violent and serious offenders, and enhances overall public safety (Casey et al., 2011).

Causality in Risk Assessment Instruments

Actuarial risk assessment tools typically use algorithms that weight different risk factors to generate a predictive risk value.[3] These tools have been used in the past to identify low-risk parole-eligible individuals, good candidates for particular programs, and those who are at highest risk of violent reoffending (Cullen & Gendreau, 2000). Risk values are used to estimate the likelihood of these particular behaviors, such as offender recidivism.[4] However, identifying and

[3] A risk factor refers to a variable that has been shown to correlate statistically with recidivism and precede recidivism in time (Kraemer et al, 1997).
[4] In the present context, recidivism is defined as “a person’s relapse into criminal behavior, often after the person receives sanctions or undergoes intervention for a previous crime” (National Institute of Justice, 2014).

operationalizing *causal* risk factors (i.e. factors that directly influence offender behavior), remains an ongoing process (Monahan & Skeem, 2014).

Risk assessment models may be "structural" or "astructural" in form. Some academics argue in favor of using astructural models (i.e. models that are purely for the purpose of *predicting* phenomena, rather than explaining *why* the phenomena occur). The argument claims that the primary goal of risk assessment is forecasting and not explanation, therefore predictors that improve forecasting accuracy should be included in the model even if the relationship between the predictor and the outcome make little intuitive sense (Berk & Cooley 1987; Berk & Bleich, 2014).[5] As such, Berk has argued that astructural models are better for short term forecasts, and structural (i.e. models based on causal relationships) with long term, as understanding the causal process is necessary for any sort of policy-relevant manipulation (Berk, 2008).

Validation

A preliminary step in the risk assessment instrument selection process includes deciding whether to develop a new tool or adopt an existing one; due to the time, money, and skills required to develop and validate a new instrument, many opt to implement a pre-existing model. For example, a tool from the same jurisdiction already in place by the probation, parole, or other supervision agency, may be modified to meet sanctioning needs, and allows for continuity, collaboration, and communication across the municipality's criminal justice system (see Latessa and Lovins, 2010). However, in the past, some jurisdictions have run into trouble implementing "off the shelf" or transplanted tools, as the model was validated on a different population, without

[5] Risk assessment instruments may be "structural" or "astructural"; the structural models attempt to represent the causal mechanisms behind recidivism, whereas astructural models utilize associations and their purely predictive (rather than explanatory) value. For example if including a particular variable improves the predictions produced by an instrument, an astructural model would include it, and regardless if the relationship between the variable and the outcome could not be explained (Berk & Cooley 1987).

considering the differences between the target population's characteristics, and which pretrial characteristics are significantly correlated with particular criminal justice outcomes, such as pretrial failure (Bechtel et al., 2011).

Additionally, testing, validating, and evaluating risk assessment instruments is of particular importance, as there are social and financial burdens that result from both false positives, and false negatives. False positives (the incorrect prediction of those at high-risk) can inadvertently increase prison populations and decrease public confidence in the practice (Berk & Bleich, 2014). However for law enforcement and the public, the cost of false negatives is believed to be higher; as Berk and colleagues learned when consulting with key stakeholders in Philadelphia, the failure to identify an offender that would be a shooter in a future offense was ten times costlier than falsely identifying an individual as a future shooter (Berk et al., 2009; Berk & Bleich, 2014).

To reduce the likelihood of false positives or negatives, initial forecasting should be implemented using three steps. First a classification scheme should be developed using pretrial predictor data that include potential predictors, and the outcome class of interest. Next the forecasting accuracy of the model should be evaluated using test data from the same jurisdiction that include the same predictors and outcome classes. If the test data forecasting is satisfactory, then the model may be used in situations when the predictors are known, but the outcome is not (Berk & Bleich, 2014). Finally, instruments should be re-validated over time to ensure that particular variables do not need to be re-weighted or otherwise adjusted as the criminal or social landscape of the jurisdiction changes, and to ensure that the instrument is correctly being implemented by well-trained staff (Andrews et al., 2006; Lowenkamp et al., 2004).

An effectiveness assessment conducted in Virginia, for example, utilized statistical tools such as Kaplan-Meier survival analysis (KM) and Cox regression. KM was used to identify individual factors significantly related to the probability of recidivism. The KM analysis could recognize which covariates were more highly correlated with rearrest across multiple time points. For example, the analysis found that significantly more males recidivated than females, as did those with prior felony drug convictions in comparison to individuals without (Kleiman et al., 2007). Additionally, Cox regression models identified predictive factors of new arrests, and new arrests resulting in conviction, when evaluating Virginia's risk assessment instrument. Unlike KM, the Cox regression evaluates an entire model, rather than individual factors (Kleiman et al., 2007).

In other aggregate validation studies, comparing different categories of instruments (e.g. unstructured, actuarial, or actuarial with dynamic factors) meta-analyses have been utilized. For example, Andrews and colleagues compared findings across existing meta-analyses of risk assessment tools based on levels of sophistication. They found that analyses of "first generation" risk instruments, which were based on unstructured clinical judgment, had relatively low predictive effect sizes ($.03 \leq r \leq .14$), whereas they were notably higher for more sophisticated, actuarial instruments ($.24 \leq r \leq .46$) (Andrews et al., 2006).[6]

It is important that validation procedures be current, and sophisticated. A simple evaluation may mistakenly attribute the results of an effective treatment for low predictive power. For example, in many states forecasts of inmate misconduct are used to determine the security level into which offenders are placed; however if higher security levels reduce the amount of inmate misconduct, the observed amount of misconduct among high-risk inmates will differ little from the amount of misconduct observed among low-risk inmates (Berk & de Leeuw, 1999). If this "suppressor effect" of higher security settings is not properly accounted for during validation, it

[6] Andrews and colleagues used Pearson's *r* to compare the predictive validity of instruments across generations; a stronger Pearson's *r* value was seen as indicative of the strength of the instrument.

can make the forecast look weak (Berk 2008). Issues such as this cause many to question the accuracy of past evaluations of criminal justice forecasts.

Principles of Risk Assessment

To maximize the impact of sentencing risk assessment in Maryland, it seems prudent to focus short term efforts on the implementation of a risk-assessment instrument (i.e. a tool to quantify individual levels of risk). The identification, verification, and evaluation of a *risk-needs* assessment (i.e. a tool to assign treatment based on risk scores) should only be considered following the successful launch of a risk-assessment instrument (which will be a sizable task in itself). As such, the implementation of a risk-needs assessment should be considered a long-term goal. The process of developing and evaluating these two types of risk instruments (risk and risk-needs) is based on three guiding principles, each described below: a risk principle, need principle, and responsivity principle.

The *risk principle*—which is of primary importance to a risk-assessment instrument—states that for the greatest impact on recidivism, the majority of services and interventions should be directed toward individuals with higher risk scores (i.e. those with a higher probability of reoffending). This idea supports the risk assessment goal of identifying high-risk offenders for specialized sanctions and/or programming. Prior research supporting the risk principle has demonstrated that many interventions are more effective on high-risk offenders' criminogenic needs (Bushway & Smith, 2007; Cullen & Gendreau, 2000). Potential explanations for this trend include: low-risk offenders are less likely to recidivate and therefore unlikely to benefit from the programming; intensive programing or supervision can potentially interrupt self-correcting behaviors already in place; treatment programs may increase exposure to high-risk offenders with pro-criminal attitudes; and treatment may disrupt pro-social networks and supports (Latessa, 2004;

Casey et al., 2011; Bushway & Smith, 2007; Cullen & Gendreau, 2000). Lowenkamp and Latessa (2004) previously identified several meta-analyses supporting the risk principle; additionally, their own research tracked over 13,000 offenders in fifty-three community-based correctional treatment facilities and revealed that the majority of programs were associated with increased recidivism for low-risk offenders and decreased recidivism for high-risk offenders.

The remaining two principles are more applicable to the long term goals of Maryland's risk based sentencing reform. The *need principle* states that correctional treatment should focus on criminogenic factors, or the needs directly linked to crime producing behavior (Casey et al., 2011). This principle argues that risk-based interventions should target dynamic risk factors (i.e. those amenable to change) associated with recidivism, as static predictors (e.g. age, race, gender) may not be manipulated. A number of studies and meta-analyses have identified three dynamic risk factors as being most predictive of recidivism: antisocial personality pattern (e.g. impulsive, adventurous pleasure seeking, restlessly aggressive and irritable), pro-criminal attitudes (e.g. rationalizations for crime, negative attitudes towards the law), and social supports for crime (e.g. criminal friends, isolation from pro-social others) (Bonta & Andrews, 2007). Other factors that are more weakly related include substance abuse, poor family/marital relationships, issues at school or work, and lack of involvement in pro-social recreational activities (Andrews & Bonta, 2006; Bonta & Andrews, 2007).

The *responsivity principle* states that the delivery of treatment programs should occur in a manner consistent with the ability and learning style of client (Bonta & Andrews, 2007). Specifically, treatment interventions should use cognitive social learning strategies and be tailored to the offenders specific learning style, motivation, and strengths (Bonta & Andrews, 2007; Casey et al., 2011). Evidence based practices and interventions are unlikely to be effective if offender characteristics (e.g. mental health conditions, level of motivation, learning style, intelligence level) are not considered when selecting an intervention (Bonta & Andrews, 2007).

Examples of Risk Assessment Tools

The instruments utilized by different jurisdictions vary in a number of ways. For example, some instruments are composed of a few, well-defined risk factors that are easily administered, while others include a broad array of risk factors that include abstract constructs and require considerable professional judgment, training, and time to implement (Monahan & Skeem, 2014). These various instruments have also been tested with varying rigor. While some proprietary "off the shelf" tools such as LSI-R (described below) have been studied many times, others have been short on empirical validation (Gottfredson & Moriarty, 2006). A meta-analysis by Yang and colleagues (2010) found that the predictive efficiencies of nine risk assessment tools were essentially interchangeable, with estimates falling within a narrow band. It is possible that instruments reach a natural limit of predictive utility, or that well validated tools may manifest similar performance because they tap common factors or shared dimensions of risk despite varied items and formats (Yang et al, 2010; Monahan and Skeen, 2014).

Additionally, despite their differences in variables and weights, some have argued that the actual statistical formulae utilized are quite similar. Generally, the predictive models are comprised of one or more dependent variables (e.g. recidivism), one or more observed predictors, and one or more unobserved disturbances (i.e. error).[7] Almost all of the forecasting done in criminal justice

[7] Models that are often used and fit this form include univariate time series, multiple (vector) time series, cross section, pooled cross-section time series, state –space, nonlinear time series, threshold autoregressive time series, autoregressive conditional heteroscedastic (ARCH), and generalized autoregressive conditional heteroscedastic (GARCH) models (Berk, 2008).

settings has historically assumed a linear relationship between predictors and outcomes; however more recent models, including predictive criminal forecasting, have begun to explore nonlinearity (Berk, 2008).

Each unique jurisdiction must select an instrument that meets their individual assessment needs, and has been properly validated for use with their unique population of offenders (Casey et al, 2011). Part of this process involves deciding whether to implement a generic “off the shelf” model, or construct a custom instrument from scratch. Examples of each are described below.

Generic Instruments

There are several “generic” risk assessment tools that are frequently utilized by parole and/or sentencing entities in the United States and Canada. These instruments are most often developed for a community supervision and corrections capacity, rather than at the sentencing decision level (Vera, 2011). Specifically, a 2010 survey conducted by the Vera Institute of Justice studying commonly used assessment tools and trends found that over sixty community supervision agencies in forty-one states reported using an actuarial assessment tool (based on responses from seventy-two agencies). Trends uncovered through the survey include: (1) assessment is relatively new (seventy percent of respondents implemented their tools since 2000), (2) state specific or state modified tools are most common (e.g. tailored versions of generic tools such as LSI-R), (3) LSI-R is the most commonly used generic tool, (4) risk *and* need are both routinely assessed, rather than just one, (5) paroling authorities generally assess risk only, (6) nearly all probation agencies conduct assessments at the pre-sentence phase to guide supervision levels, (7) sharing results across agencies is common, and (8) storage of results is nearly all electronic.

The generic tools utilized in many states (in addition to some Canadian provinces) include the Level of Service Inventory-Revised (LSI-R), Level of Service/Case Management Inventory (LS/CMI), Correctional Offender Management Profiling for Alternative Sanctions (COMPAS), and Static-99.

The LSI-R is described as a "quantitative survey of offender attributes and their situations relevant to level of supervision and treatment decisions." The instrument, which is designed for offenders aged sixteen and up, helps to make security level classifications, decisions about probation and placement, and assess treatment progress. The fifty-four item scale is comprised of ten sub-scales: criminal history, education/employment, financial, family/marital, accommodation, leisure/recreation, companion, alcohol/drug problem, emotional/personal, and attitude/orientation scales (Andrews & Bonta, 1995). The LS/CMI tool is an updated version of the original LSI-R, which also serves as a case management device by incorporating supplementary information. The tool condenses the 54 LSI-R items into 43 items, and adds ten additional sections, with the first section (the LSI-R) able to operate as a stand-alone tool (Andrews et al., 2004). A 2004 meta-analysis revealed the instrument's predictive value, and demonstrated that the LS/CMI is as predictive and reliable with females as it is males (Williams et al., 2009). Additionally, a Screening Version (LSI-R: SV) can be administered when issues of resource and time constraints exist. The LSI-R: SV is a parsed down, eight item tool that may be used to determine whether the full LSI-R needs to be run (it does not serve as a stand-alone tool).

Although the LSI-R was developed primarily as a supervision and security level tool, it has also been applied at the pre-sentence level in numerous jurisdictions. For example, in Tulsa, Oklahoma offenders eligible for community sanctions may receive the LSI-R in order to determine eligibility for community sentencing.[8] Similarly, in the Fifth District of Idaho, the LSI-R is completed at the presentence stage, with results included in the Presentence Investigation Report.[9]

The COMPAS assessment instrument is used to identify program needs regarding placement, supervision, and case planning. It consists of a computerized database and analysis program that provides separate risk estimates for future risk for violence, recidivism, failure to appear, and community non-compliance (i.e. technical violations). COMPAS also includes a "criminogenic and needs profile" for offenders, which includes information about the individual's criminal history, needs assessment, criminal attitudes, social environment, and social support. The assessment domains include both risk and protective factors, which include job and educational skills, finances, bonds, social support, and non-criminal parents and friends. Generally, evidence has been supportive of COMPAS, with a recent study indicating that the predictive validity of the instrument matches or exceeds other actuarial tools, and performs equally well in predicting across race and gender (Brennan et al., 2009; Vera, 2011).

Similar to LSI-R, while COMPAS was developed for the corrections stage of the criminal justice process, numerous jurisdictions have begun implementing it during the pre-sentence investigation phase. For example, nearly thirty probation departments in New York State are currently actively using COMPAS and are preparing to develop a system to integrate COMPAS results with portions of the pre-sentence investigation.[10]

Lastly, the Static-99 is the most commonly used actuarial risk assessment tool for predicting sexual and violent recidivism among adult male sex offenders (Hanson & Thornton,

[8] http://www.tulsacounty.org/communitysentencing/
[9] http://fifthdcs.com/FifthPolicy/index.cfm?policy=PsiIntroduction
[10] http://www.criminaljustice.ny.gov/opca/technology.htm

2000). The tool contains ten items: age less than 25, never lived with a lover for two years, any prior convictions for non-sexual violence, any current convictions for non-sexual violence, four or more prior sentencing dates, prior sexual offenses, non-contact sexual offenses, any male victims, any unrelated victims, and any stranger victims (Harris et al, 2003). The Static-99 is used for treatment planning, community supervision, and civil commitment evaluations for this specific type of offender (Helmus et al., 2012). The predictive accuracy of the Static-99 has been assessed at least sixty-three times; studies have found high levels of rater reliability (Harris et al., 2003), moderate accuracy in predicting sexual recidivism risk (Hanson and Morton-Bourgon, 2004), but large and significant variability across studies (Helmus et al., 2012).

Static-99 is utilized at the pre-sentence phase in some jurisdictions, such as in California. Specifically, California uses the risk score to determine levels of supervision for offenders on probation, parole, or forms of intermediate sanctions.[11] Similarly, the Static-99 score is included in the specialized sex offender Pre-Sentence Investigation Report that a judge receives prior to imposing a sentence.[12]

Non-Generic and Jurisdiction-Specific Instruments

In contrast to generic "off-the-shelf" instruments, other jurisdictions have chosen to develop actuarial instruments from the ground up, typically with the assistance of a team of researchers. These include the use of individualized advanced statistical models, and are developed and validated using the specific population they intend to forecast.

[11] http://www.meganslaw.ca.gov/riskfaq.aspx?lang=ENGLISH
[12] http://doc.vermont.gov/about/policies/rpd/correctional-services-301-550/335-350-district-offices-general/copy_of_342-01-pre-sentence-investigation-psi-reports

Some academics have advocated for jurisdictions to use classification and regression tree (CART) based machine learning models.[13] Classification trees construct profiles of individuals associated with different outcome classes, by breaking the data into groups with similar profiles. Proponents of these models assert that while they are high in accuracy, they are quite unstable, as the introduction of new data generates different profiles (Berk & Bleich, 2014). Some CART models have been shown to provide more reliable predictions than other models. For example, random forests predictions (CART models with an additional layer of randomization) have been shown to: significantly control over-fitting, inductively determine nonlinear relationships, allow highly specialized predictors to participate, and allow for asymmetric costs of false positives and false negatives (Berk 2008). Additionally, random forests models address poor predictors by sampling predictors as each classification or regression tree is grown.

The random forests approach was evaluated when used to predict murderous conduct (charges of homicide or attempted homicide) by offenders on probation and parole in Philadelphia over a two-year period (Berk et al., 2009). Overall, ninety-four percent of individuals who did not commit a homicide or attempted homicide were correctly forecasted by the random forests model. Forty percent of the individuals who committed a homicide or attempted homicide were correctly forecasted (Berk et al., 2009).[14]

[13] Examples of CART models include random forests, Bayesian additive regression trees, and stochastic gradient boosting.

[14] The astructural model tested restricted the assessment of risk to violent recidivism, and built in the relative costs of false positives and negatives. The outcome of homicide or attempted homicide was rare by most any standard, as only about one percent of individuals under supervision would be expected to fail by this definition, making forecasting particularly demanding. They found that the most significant predictors in their model are age of the offender, age of first contact with the adult court system, and number of prior convictions involving a firearm. Additionally, they also found that the interaction between certain predictors, namely age, age at first contact with the adult court system, and number of prior violent convictions, and the probability of being charged with homicide or attempted homicide are highly non-linear (Berk et al, 2009).

Models such as random forests are best implemented by computers linked to large criminal justice databases—this is a capacity that does not currently exist in many courtrooms.[15] As such, Berk and colleagues have recently begun to advocate for a simpler classification tree approach to forecasting, which overcomes some of the machine learning and IT infrastructure requirements that make random forests impractical for many jurisdictions (Berk and Bleich, 2014). These classification tree models include the profiles of individuals associated with different outcome classes, but eliminate the additional randomization component involved in random forests.

While approaches such as CART models and generic instruments are common, many of the states that utilize risk assessment tools at the sentencing level have developed or adapted their own unique approach. The number of states that have utilized risk assessment tools to inform sentencing, relative to the number using them at other points in the criminal justice process, is rather limited. The states with the clearest examples of integrating risk assessment into sentencing decisions are Virginia, Pennsylvania, Missouri, and Utah. Other states such as Ohio and Georgia use risk assessment at various stages, but do not appear to have an explicit sentencing risk assessment program.

State sentencing instruments

Among the states that use risk assessment in sentencing, a variety of tools are currently being employed (see Table 1, p.23). For example, the Virginia State Criminal Sentencing Commission was mandated by legislation to produce an empirically based risk assessment

[15] Maryland's Criminal Justice Information System (CJIS) Central Repository receives, maintains, and disseminates Maryland's criminal history records. It receives reports of criminal "events" (e.g., arrests, convictions, sentences, etc.) from law enforcement, courts, corrections, and other criminal justice entities. Depending on the restrictiveness, and nature of this database, random forests assessment may in fact be more practical for the state of Maryland than many other jurisdictions.
http://www.dpscs.state.md.us/publicinfo/publications/pdfs/CJIS/CJIS_CR_Primary_Services.pdf

instrument (VA. Code Ann. § 17-235(5) (1995)). The state's Commission developed its own instrument, "Worksheet D", to identify low-risk offenders that would make suitable diversion candidates. This tool is used to divert "25% of the lowest risk, incarceration-bound, drug and property offenders for placement in alternative (non-prison) sanctions" with the goal of decreasing the prison population without creating a significant risk to public safety (Kern & Farrar-Owens, 2004; Soulé & Najaka, 2013). If an offenders calculated risk score is below a set threshold, the otherwise incarceration bound offender is recommended for alternative sanctions.[16]

Similar to Virginia, Pennsylvania's adoption of sentencing risk assessment was guided by 2009 legislation that a required an evidence-based, validated risk assessment tool to be used to reduce recidivism and public safety risks, and to maximize reentry success. The legislation indicated that the guidelines should adopt a risk assessment tool to be used at sentencing, consider the risk of re-offense and threat to public safety, help determine if the offender is eligible for alternative sentencing programs, and develop an empirically based worksheet using factors predicting recidivism. Pennsylvania's instrument development process is ongoing; at the present time the Pennsylvania Sentencing Commission is testing and evaluating the strength of various risk factors, and communications methods, which will both impact the structure of the final model.

Missouri utilizes an indeterminate sentencing structure, but incorporates risk assessments into presentence reports provided to judges, ensuring judges are "fully informed" when using discretion (Hyatt et al, 2011; Wolff, 2006). Sentencing Assessment Reports, which include the Commission's sentencing recommendations (based upon the guidelines grid), also include information regarding the offense, offender risk factors, a management plan, the Commission's

[16] The instrument is only scored for offenders who meet the sentencing guidelines recommendation for incarceration (probation cases are not considered for diversion), and a criminal history of only nonviolent offenses (i.e. larceny, fraud, and drug offenders). Current or prior violent felony convictions, and selling an ounce or more of cocaine excludes one from risk assessment consideration.

recommendation, and the Parole Board guidelines for release when a prison sentence results. The Sentencing Assessment Report and an online Automated Recommended Sentencing Information application (www.mosac.mo.gov) calculate the offender risk assessment score. The score is publicly available on the website so that prosecutors and defense attorneys may incorporate it into plea negotiations.

Utah uses the LSI-R for all convicted felons, and incorporates the results of the assessment into a presentence investigation conducted by the Department of Rehabilitation and Corrections (Monahan and Skeem 2014). When imposing a sentence, the judge must consider a sentence calculated under guidelines in addition to an LSI-R influenced recommendation. Utah also uses the calculated LSI-R score to assist in determining conditions of one's probation, including the content of treatment offered and level of supervision.

Factors utilized for prediction by states

During development and pilot testing of Virginia's Worksheet D, the Commission identified eleven statistically significant factors in predicting recidivism and assigned scores based on their relative importance (each is scored separately, with the sum providing the overall risk score). Risk factors include: offender age, gender, marital status, employment status, whether acted alone, whether additional offenses at conviction, whether arrested or confined within past twelve months, prior criminal record, prior drug felony convictions, if incarcerated as an adult, and if incarcerated as a juvenile. Specifically, the variables age, prior record, and prior juvenile incarceration are most heavily weighted; however a 2007 analysis found that while the tool is effective, it requires some re-weighting, including weighting prior record and gender more heavily (Kleiman et al, 2007).

Pennsylvania based recommendations for risk factors on existing instruments and the Pre-Sentence Investigation reports; the state's Commission recommended that the model include both static and dynamic risk factors, conduct future research (to evaluate the extent risk information is currently collected and the ability to use it to predict recidivism) and to modify the Sentencing Guidelines Software Web to include questions related to validated risk factors. In the ongoing development process, the factors that best predicted recidivism for solitary conviction offenders were also the best predictors for multiple conviction offenders. The eight factors that best predicted recidivism—gender, age, county, total number of prior arrests, prior property offenses, prior drug arrests, whether they are a property offender, and an offense gravity score—generate a risk score ranging from 0-14, which is grouped into two risk categories (low and high-risk).

The characteristics included in Missouri's risk assessment tool are classified as "offense-related factors" or "other risk-related factors". Offense related factors include variables such as prior unrelated findings of misdemeanor or felony guilt, prior incarceration, five year periods without guilt or incarceration, revocations of probation or parole, current offense recidivism-related; other risk related factors include age, prior escape, education, and employment. The individual risk factors are assigned values ranging from -1 to +2, and are summed across eleven fields; a total score of 4 to 7 is rated "good, 2 to 3 is "above average", 0 to 1 is "average", -1 to -2 is "below average", and -3 to -8 is "poor". These scores are incorporated into the Sentencing Assessment Report.

Lastly, Utah utilizes the variables indicated by the LSI-R instrument. Utah stakeholders defended this decision by arguing that no tools seemed to reliably outperform the Level of Service instruments, and that the new tool should be, equally predictive and valid, have lower costs related to training and adoption, and to increase the agency's economic efficiency (Prince & Butters, 2014).

Table 1: Cross State Comparison of Risk Assessment Instruments[17]

State	Virginia		Pennsylvania	Missouri	Utah	Georgia
Tool	**Nonviolent Offender Risk Assessment**	**Sex Offender Risk Assessment**	**Risk/Needs Assessment**	**Offender Risk Assessment**	**LSI-R**	**COMPAS**
Number of Items	11	9	8	11	54	Varies
Domains	• Offender characteristics and demographics • Current offense information • Prior adult criminal record • Prior juvenile record	• Offender age • Prior person/sex arrests • Offender relationship and victim age • Employment status • Offense location • Prior sex offender treatment • Prior incarcerations • Education • Nature of sex offense	• Gender • Age • County • Prior arrests • Prior drug offenses • Prior property offenses • Property offender • Offense gravity score	• Criminal history • Age • Prior escape • Education • Employment	• Criminal history • Education/employment • Financial • Family/marital • Accommodation • Leisure/recreation • Companions • Alcohol/drug problem • Emotional/personal • Attitudes/orientation	• Criminal history • History noncompliance • History of violence • Current violence • Criminal associates • Substance abuse • Financial problems • Vocational/educational • Criminal attitudes • Family criminality • Social environment • Leisure • Residential instability • Criminal personality • Social Isolation
Application	Sentencing Decisions	Sentencing Decisions	Sentencing Decisions	Sentencing Decisions and Plea Negotiations	Sentencing Decisions	Programming and Reentry

State	Ohio				
Tool	**ORAS-CSST**	**ORAS-CST**	**ORAS-PAT**	**ORAS-PIT**	**ORAS-RT**
Number of Items	4	35	7	31	20
Domains	• Number of prior felonies • Current employment • Availability of drugs • Number of criminal friends	• Criminal history • Education • Employment/finances • Family/social support • Neighborhood problems • Substance abuse • Antisocial associations • Antisocial attitudes/behavioral problems	• Criminal history • Employment • Residential stability • Substance abuse	• Age • Criminal history • Education • Employment/financ es • Family/social support • Substance abuse • Criminal lifestyle	• Age • Criminal history • Social bonds • Criminal attitudes
Application	Community Supervision Screening	Community Supervision	Pretrial Assessment	Prison Intake	Reentry

[17] This table was adapted from a similar table presented in the Vera Institute of Justice's (2011) report to the Delaware Justice Reinvestment Task Force

References

Ægisdóttir,S. While, M.J., Spengler, P.M., Maugherman, A.S., Anderson, L.A., Cook, R.S. et al. 2006. The meta0analysis of clinical judgment project: Fifty-six years of accumulated research on clinical versus statistical prediction. *The Counseling Psychologist, 34,* 341-382.

Andrews, D. A., Bonta, J., and Wormith, S. J. 2004. *The Level of Service/Case Management Inventory (LS/CMI)*. Toronto: Multi-Health Systems.

Andrews, D. A., Bonta, J., and Wormith, S. J. 2006.The recent past and near future of risk and/or need assessment. *Crime and Delinquency, 52*, 7-27.

Andrews, D. A., and Bonta, J. 2006. *The psychology of criminal conduct* (4th ed.). Newark, NJ: LexisNexis.

Bechtel, K., Lowenkamp, C., and Holsinger, A. 2011. Identifying the predictors of pretrial failure: A meta-analysis. Final Report.

Berk, Richard A. 2008. Forecasting methods in Crime and Justice. *Annual Review of Law and Social Science (Annual Reviews Press)* 4, 219–238.

Berk, R., and Bleich, J. 2014. Forecasts of violence to inform sentencing decisions. *Journal of Quantitative Criminology, 30*(1), 79-96.

Berk, R. A., and Cooley, T. F. 1987. Errors in forecasting social phenomena. *Forecasting In The Social And Natural Sciences* (pp. 247-265). Springer: Netherlands.

Berk, Richard A., and de Leeuw, J. 1998. An evaluation of California's inmate classification system using a generalized regression discontinutity design. *Journal of the American Statistical Association, 94,* 1045-52.

Berk, R.A., Sherman, L., Barnes, G., Kurtz, E., and Ahlman, L. 2009. Forecasting murder within a population of probationers and parolees: a high stakes application of statistical learning. *Journal of the Royal Statistical Society: Series A (Statistics in Society), 172(1), 191-211.*

Bonta, J., and Andrews, D. A. 2007. Risk-need-responsivity model for offender assessment and rehabilitation. *Rehabilitation, 6.*

Brennan, T., Dieterich, W., and Ehret, B. 2009. Evaluating the predictive validity of the COMPAS risk and needs assessment system. *Criminal Justice and Behavior,* 36, 21-40.

Bushway, S., and Smith, J. 2007. Sentencing using statistical treatment rules: what we don't know can hurt us. *Journal of Quantitative Criminology, 23*(4), 377-387.

Casey, P. M., Elek, J. K., Warren, R. K., Cheesman, F., Kleiman, M., and Ostrom, B. 2014. Offender risk and needs assessment instruments: A primer for courts. Williamsburg, VA: National Center for State Courts. Retrieved from http://www.ncsc.org/~/media/Microsites/Files/CSI/BJA%20RNA%20Final%20Report_Combined%20Files%208-22-14.ashx.

Champion. D.J. 1994. *Measuring offender risk: A criminal justice sourcebook.* Westport: Greenwood Press.

Cullen, F.T. and Gendreau, P. 2000. Assessing correctional rehabilitation: Policy, practice, and prospects. *Policies, Processes, and Decisions of the Criminal Justice System*, 3, *109-175.*

Goldkamp, J., and Gottfredson, M. 1985. Policy Guidelines for Bail: An Experiment in Court Reform. Philadelphia: Temple University Press.

Gottfredson, S., and Moriarty, L. (2006). Statistical risk assessment: Old problems and new applications. *Crime & Delinquency, 52(1),* 178-200.

Grove, W. M., Zald, D. H., Lebow, B. S., Snitz, B. E., and Nelson, C. 2000. Clinical versus mechanical prediction: A meta-analysis. *Psychological Assessment, 12*, 19-30.

Hanson, R. K., and Thornton, D. (2000). Improving risk assessments for sex offenders: A comparison of three actuarial scales. *Law and Human Behavior, 24(1),* 119-136.

Hanson, R.K. and Morton-Bourgon, K. 2004. Predictors of sexual recidivism: An updated meta-analysis. *Public Works and Government Services Canada.*

Harris, G.T., Rice, M.E., Quinsey, V.L., Lalumiere, M.L., Boer. D., and Lang, C. 2003. A multi-site comparison of actuarial risk instruments for sex offenders. *Psychological Assessment, 15*, 413-25.

Heilbrun, K. 2009. *Evaluation for Risk of Violence in Adults.* New York: Oxford University Press.

Helmus, L., Thornton, D., Hanson, R.K. and Babchishin, K.M. 2011. Improving the predictive accuracy of Static-99 and Static-2002 with older sex offenders: Revised age weights. *Sexual Abuse: A Journal of Research and Treatment*, *24,* 64-101.

Hyatt, J.M., Chanenson, S.L. and Bergstrom, M.H. 2011. Reform in motion: The promise and perils of incorporating risk assessments and cost-benefit analysis into Pennsylvania sentencing. *Villanova Public Law and Legal Theory Working Paper Series.*

Kleiman, M., Ostrom, B.J., and Cheesman, F.L. 2007. Using risk assessment to inform sentencing decisions for nonviolent offenders in Virginia. *Crime & Delinquency, 53,* 106-32.

Kraemer, H.C., Kazdin, A.E., and Offord, D.R. 1997. Coming to terms with terms of risk. Archives of General Psychiatry, 54, 337-43.

Latessa, Edward J. (2004). The challenge of change: Correctional programs and evidence-based practices. *Criminology & Public Policy, 3(4)* 547-60.

Latessa, E.J. and Lovins, B. 2010. The role of offender risk assessment: A policy maker guide. *Victims & Offenders, 5(3)* 203-19.

Lowenkamp, C. T., Latessa, E. J., & Holsinger, A. M. 2004. Empirical evidence on the importance of training and experience in using the Level of Service Inventory-Revised. *Topics in Community Corrections*, 49-53.

Monahan, J., & Skeem, J. L. (2014). Risk Redux: The Resurgence of Risk Assessment in Criminal Sanctioning. *Federal Sentencing Reporter*, *26*(3), 158-166.

Palacios, V. J. (1994). Go and Sin No More: Rationality and Release Decisions by Parole Boards. *South Carolina Law Review, 45*:567-615.

Prince, K., and Butters, R. P. 2014. Brief Report: An Implementation Evaluation of the LSI-R as a Recidivism Risk Assessment Tool in Utah.

Silver, E., and Miller, L. L. 2002. A cautionary note on the use of actuarial risk assessment tools for social control. *Crime & Delinquency*, *48*(1), 138-161.

Soulé, D.A. and Najaka, S.S. 2013. Paving a path to informed sentencing decisions. *Federal Sentencing Reporter, 25(3)* 181–184.

VanNostrand, M. and Lowenkamp, C.T. 2013. Assessing pretrial risk without a defendant interview. *Laura and John Arnold Foundation*. Retrieved from: http://www.arnoldfoundation.org/sites/default/files/pdf/LJAF_Report_no-interview_FNL.pdf

Vera Institute of Justice. 2011. Risk and Needs Assessments. Retrieved from: http://ltgov.delaware.gov/taskforces/djrtf/DJRTF_Risk_Assessment_Memo.pdf

Williams, K. , Andrews, D. , Bonta, J. , Wormith, J. , Guzzo, L. and Brews, A. (2009, March 4) "The Level of Service/Case Management Inventory (LS/CMI): Reliability and Validity in Female Offenders" Paper presented at the annual meeting of the American Psychology - Law Society, TBA, San Antonio, TX

Wolff, J. 2006. Risk, fear, blame, shame and the regulation of public safety. *Economics and Philosophy*, *22*: 409-27.

Yang, M., Liu, Y., and Cold, J. 2010. Applying neural networks and other statistics models classification of serious offenders and the prediction of recidivism. Vol. 6/10. London: Ministry of Justice.