Are Intimate Partner Violence (IPV) risk assessment tools racially biased? Some observations from research and practice

Kathleen J. Ferraro, Ph.D.
Neil S. Websdale, Ph.D.*
Family Violence Institute
Northern Arizona University
Flagstaff, Arizona

*Denotes equal authorship. Correspondence to Family Violence Institute, Northern Arizona University, PO Box 15026, Flagstaff, AZ. 86011-15026.

The Battered Women's Justice Project
1801 Nicollet Ave South, Suite 102, Minneapolis MN 55403
technicalassistance@bwjp.org 800-903-0111, prompt 1

Funding for this document was made available through the US Department of Health and Human Services, Grant #90EV0440. The viewpoints contained in this document are solely the responsibility of the authors and do not represent the official views or policies of the department and do not in any way constitute an endorsement by the Department of Health and Human Services.

Introduction

On November 20, 2016, King Yates shot and killed his wife, 24-year-old Cassandra Yates. King had multiple prior arrests for domestic violence against a former girlfriend and against Cassandra. About a year before the murder, in jail awaiting trial on drug and firearm charges, he was diagnosed with Bi-Polar disorder and prescribed anti-psychotic medication. He was found not guilty and released. In May 2015, he was arrested for stabbing Cassandra in the neck, but she changed her original story and claimed the wound was the result of an accident. No charges were filed. Two weeks before the murder, King was arrested for prohibited possession of a firearm. He was out of jail in three days, released on $1,100 bond, which Cassandra paid.[1]

The day after Cassandra's murder, King was booked into jail on first-degree murder charges where he shared a cell with 24-year-old Brandon Roth. Brandon was awaiting sentencing for a nonviolent property offense. In April 2016, Brandon was found dead in their cell after King beat him with a sock filled with batteries and strangled him with a cord.[2]

These murders illustrate two ways that risk assessments within the criminal justice system might have led to enhanced safety for victims and decreased incarceration of low risk offenders. On the one hand, had King Yates been assessed for risk of future serious or lethal intimate partner violence (IPV), he would have scored as "high risk" on any of the existing tools. He may have been held longer in jail and provided more intensive supervision on release. King may not have been able to acquire another firearm and may not have murdered Cassandra. If Cassandra had been linked to victim services as a result of completing an IPV risk assessment (RA), she may have been more cognizant of the danger

she faced and able to escape King's control. On the other hand, if a pre-sentence offender RA determined Brandon Roth was low risk and he had been released on his own recognizance, he would not have been locked in a cell with a murderer and would not have been brutally killed. A high risk, dangerous intimate partner offender would have been detained, potentially saving the life of his victim and a low risk property offender would have been released, saving him from murder in his cell.

For many practitioners and researchers, RAs provide relevant, useful guidance on a range of legal decisions regarding case management. As one component of the trend toward "evidence-based practice," many argue these tools offer objective, scientific knowledge about "what works" with whom. But others see them as potentially discriminatory, disproportionately identifying certain racial and ethnic minorities as being at "high" or "elevated" risk of re-offending.

In what follows, we consider the question of whether IPV RAs are racially biased. We commence with an overview of the way RAs proliferated as part of the rise of what sociologists and criminologists refer to as risk societies[3] and cultures of control,[4] social arrangements that, among many other characteristics, evidence deep patterns of racial, class, and gender inequality. We note the links between RA tools, bail reform, and changes in pre-trial detention. Next, we introduce case law that cautions the assessment of risk might involve discrimination against groups such as certain racial and ethnic minorities. We overview the research studies that contend there is racial bias in assessing risk and examine those studies that counter such claims. The discourse on risk and racial bias raises important issues about ethics and fairness. We explore some of these issues.

Our discussion of the discourse on RA lays the groundwork for our specific consideration of whether IPV RAs are racially biased. We start by describing the development of IPV RAs amidst growing research and practice in the field of violence against women. We discuss how these tools work and some of the pros and cons of using them. Our discussion leads us to ask whether IPV RAs might be racially biased against the accused, the alleged victim, or both parties. Importantly, the extant research does not answer the extraordinarily complex question of whether IPV RAs are racially biased. We therefore stress the need to consider the implications of using IPV RAs in the broader movement to confront IPV, particularly noting their limitations and tendency to mask bigger questions.

Risk societies and the culture of control in the United States

The contemporary United States serves as an example of a risk society. Many people have become obsessed with assessing risks associated with blood cholesterol, hazardous weather systems, global warming, and, for our purposes, the likelihood of intimate partner homicide (IPH). With the erosion of the seeming sureties of pre-modern tradition, religion, ritual, ceremony, and custom, our desire for predicting future adverse outcomes has greatly increased. Intensifying concerns about "risk" reflect growing needs for certainty in what many see as fast-changing, uncertain, anxious times. Using specialist knowledge to assess and possibly lower the risk of adverse outcomes is a way of trying to know or colonize the future, to regulate it, and render it less threatening, thus increasing our security in the face of increasing perceptions of insecurity.

IPV risk assessments can create a shared language of "danger" among the community agencies and stakeholders charged with dealing with these cases. Their

growing use since the 1980s reflects the increase in general fears about crime and security since the late 1970s. These concerns are ironic given falling crime rates, general improvements in life expectancy, and, for our purposes, declining rates of IPV and IPH.[5]

Criminologist David Garland sees these and other developments as part of an expanding "culture of control" emblematic of political and cultural life in the post 1970s U.S and U.K. Garland identifies growing networks of criminal justice and community partnerships and interagency agreements designed to enhance crime prevention and improve community safety. These regulatory networks blossomed alongside mass incarceration, globalization, the technological displacement of labor, the shrinkage of state infrastructures, and the increasing polarization of wealth, exacting a heavy toll on the poor and certain racial and ethnic minority communities. The growth of private security, gated communities, the private prison sector, state surveillance, and the commercialization of crime control, all invite questions such as: Who benefits and who suffers from these developments? Pointedly, Garland suggests that our very notions of "danger" target street offenders and disadvantage minorities in urban centers and blighted rural communities. For Garland, these are political foci, perhaps distracting attention from much bigger problems such as corporate crime, white-collar crime, and environmental pollution. Significantly, he identifies an accelerating "economic style of decision making" that mirrors an increasingly deregulated economy, and a growing hostility to providing for the disadvantaged, whether they be homeless people, low wage and increasingly non-unionized workers, or the mentally ill.

We must also note what Garland refers to as the "remarkable return of the victim to center-stage in criminal justice policy."[6] In colonial and pre-revolutionary times crime

victims used to initiate prosecutions and actively participate in the criminal justice process. Their involvement waned only to pick up again toward the end of the 20th century, spurred on in the US by the Victim Rights Act (2005). Amidst these developments Garland notes a growing interest in the feelings of victims and their families. He notes, "The new political imperative is that victims must be protected, their voices must be heard, their memory honored, their anger expressed, their fears addressed."[7] Relatedly, Garland opines, "Any untoward attention to the rights or welfare of the offender is taken to detract from the appropriate measure of respect for victims".[8] Increasingly, RAs in general and IPV RAs in particular, examine the lives, voices, and traumas of victims, rendering some of this sensitive personal information available to the criminal and civil justice systems. Those raising concerns about the tools undermining the civil liberties and due process rights of the accused, have indeed faced criticisms that they might be disrespecting victims.

The development of RA tools

Criminal justice professionals, from law enforcement to probation officers to judges, have used available information to determine offenders' risk since the early 1900s. They have employed knowledge of prior criminal activities, family and occupational status, substance abuse and mental health to inform decisions about pre-trial release, supervisory status, sentencing and parole.[9] Beginning in the 1930s, RAs were developed based on mathematical formulae, algorithms, which utilized data on thousands of criminal offenders to predict the likelihood of future crime. Known as actuarial RAs, these tools are in widespread use for the general criminal population. Their use accelerated in the late 20th century and contributed to the habitual offender and three-strike laws that sought the incapacitation of offenders posing the greatest threat to public safety. During this time,

prior criminal record became the most salient and reliably predictive factor in assessing future risk, an issue that we explore further in our discussion of racial bias.

By 2018, RAs now inform decision-making regarding pre-trial services, probation, and sentencing. Many favor their adoption as an objective approach to reducing jail and prison populations and minimizing unfairness by replacing subjective, often unacknowledged, biases with mathematically derived predictions.[10] The majority of offenders are classified as low risk for flight and further criminal activity while awaiting trial or sentencing, thus reducing the need for their detention. Recently developed RAs adopt the Risk-Needs-Responsivity (R-N-R) model to craft treatment interventions designed to prevent recidivism.[11] For example, non-violent offenders assessed to be substance abusers might receive drug treatment instead of remaining in jail. Proponents argue such "risk-needs" assessments provide empirically justified grounds for restoring rehabilitative correctional goals and reducing reliance on incarceration to manage risk.[12]

Critics of RA tools raise concerns about contributing to faster, assembly line justice that relies on tools that may unfairly punish offenders, misrepresent treatment needs, and contribute to existing social inequality. Some note the injustice of relying on statistical predictions of future harm rather than the crimes that brought offenders into the system.[13] Numerous critics argue that RAs act as a mere triaging mechanism for an overworked, under resourced criminal and civil justice system, in a way perpetuating the system. For them, the focus on potential recidivism diverts attention from the contexts that lead people to engage in criminal conduct, such as high rates of poverty and institutionalized racism.[14] Among other negative results of RA, these critics draw attention to the ways RAs reinforce

and exacerbate existing racial, gender and economic disparities.[15] We review the research on these critiques below.

Bail reform and pre-trial detention

Following arrest, offenders attend an initial appearance where an officer of the court decides whether to release them on their own recognizance or detain them in jail. Those detained may have a monetary amount set, known as bail, to secure their appearance at future hearings. The Constitution requires that pre-trial detention be rare and fully justifiable, but the 1984 Bail Reform Act ruled that there is no Constitutional prohibition on detention of offenders who pose a risk to the community.[16] The United States has one of the highest pre-trial detention rates in the world and the majority of those in jail are unable to meet the costs of bail, even at very low levels.[17] Pre-trial detention is also an expensive intervention, costing taxpayers $14 billion each year. As a consequence, there is a national movement to eliminate bail and limit detention only to those at greatest risk of flight or recidivism. This serves the interests of justice by eliminating the economic unfairness generated by money bail, the negative consequences of detention, and the high costs of incarceration, while preserving public safety.

IPV offenders constitute a particularly problematic category in terms of determining conditions of release at an initial hearing or arraignment.[18] For misdemeanor offenses, which comprise the majority of IPV charges, these hearings are sometimes presided over by magistrates who may have limited legal education and scant information about an offender prior to determining whether to detain with bail, the amount of bail, or to release on their own recognizance.

Many jurisdictions now include results of a RA at these initial hearings. Some states statutorily define the types of information judicial officers must consider, and these increasingly include IPV RAs. For example, Arizona Revised Statute 13-3967 B (5) states the judicial officer shall take into account, "the results of a risk or lethality assessment in a domestic violence charge presented to the court." However, the judicial officer must also take into account 14 other issues including some that the risk or lethality assessment might address. For example, one IPV RA, the Arizona intimate Partner Risk Assessment Instrument System [APRAIS] asks victims whether they think the accused is capable of killing them. A RA may also contribute information that assists judicial officers in determining the danger the accused presents to others in the community and the views of the victim (e.g., the APRAIS broadly canvasses victims on these matters).

These RA tools are controversial for many, often interrelated, reasons. For example, the APRAIS tool would have identified King Yates, an African American man, as being at high risk of committing severe re-assault within seven months from the time it was initially administered. We have already discussed how the administration of such a tool might have proved useful. But critics of RA tools might well ask how many other young, African-American males, screened as "high-risk" by the APRAIS tool, might have been denied bail based on the accusations of "alleged" victims. These critics cite the damaging effects of spending even a short period of time in pre-trial detention and the possibility that African-American males might be disproportionately subject to such possibly unfair confinement. The courts have expressed similar concerns in relation to RAs. It is in that direction we now turn.

Case law and cautions about possible racial bias in assessing risk

As the courts are now beginning to note, some proprietary risk assessments have not been subject to the gold standard of blind, independent peer review by researchers. Neither are their proprietary algorithms subject to public scrutiny. Effectively, the algorithms are trade secrets. These render them difficult to challenge and at the same time confront the question of whether they might be racially biased. The Wisconsin Supreme Court recently addressed these issues in the *Loomis* case.[19] The rulings of the court speak directly to the question of racial bias and risk assessment. Notably, the United States Supreme Court declined to hear Eric Loomis's petition for Writ of Certiorari, after the Wisconsin Supreme Court affirmed the decision of the trial court.[20]

Eric Loomis was charged with five criminal counts related to a drive by shooting. He admitted to driving the vehicle but denied participating in the shooting. He eventually pleaded guilty to two charges, "attempting to flee a traffic officer and operating a motor vehicle without the owner's consent."[21] The presentence investigation report (PSI) contained reference to the COMPAS RA,[22] a closed-source RA software. The trial court considered the COMPAS RA in its sentencing of Loomis to six years of imprisonment and five years of extended supervision. Appealing the sentencing decision, Loomis argued the use of the COMPAS tool violated his due process rights.[23] Essentially, Loomis claimed the COMPAS tool relied on data from groups of people and that its proprietary nature meant he could not know the basis for the assessment in his individual case. Since the COMPAS tool used "gender" in its algorithm, he claimed his due process rights were further breached because such a consideration of gender was "unconstitutional."

The Wisconsin Supreme Court affirmed the decision of the lower court, rejecting Loomis's due process arguments.[24] According to the Wisconsin Supreme Court, the use of gender, "served the nondiscriminatory purpose of promoting accuracy."[25] Importantly, the Court held that Loomis had not provided sufficient evidence that the sentencing court had considered gender. For purposes of our question, Loomis's failure to provide evidence of the sentencing court's consideration of gender raises important questions about how those alleging race discrimination in the use of IPV RAs can sufficiently demonstrate whether a court took "race" or a possible proxy measure of "race" into consideration.

The Wisconsin Supreme Court issued a number of cautions regarding the use of risk assessment tools. The court warned that such tools cannot be used, "to determine whether an offender is incarcerated" or "to determine the severity of the sentence."[26] Importantly, PSIs that use a COMPAS assessment must include five written warnings for judges, specifically that:

1) The proprietary nature of the tool prevents disclosure of how risk scores were determined.

2) Scores are unable to identify particular high-risk individuals because they derive from group or population data.

3) The COMPAS algorithm is based on national datasets not from data specific to the state of Wisconsin. Put differently, COMPAS has not been validated for the Wisconsin population.

4) "Studies have raised questions about whether [COMPAS scores] disproportionately classify minority offenders as having a higher risk of recidivism."[27] Here the court sought to instill a "targeted skepticism" regarding the matter of possible racial discrimination.

5) COMPAS was developed to help the Wisconsin Department of Corrections make post-sentencing determinations.

A Harvard Law Review summary of *Loomis* notes the Wisconsin Supreme Court's attempt to instill a "general skepticism about the tool's accuracy and a more targeted skepticism with regard to the tool's assessment of risks posed by minority offenders."[28] However, the summary concludes that the court failed to state the force of the criticisms regarding the possibility of disproportionately classifying minorities as high-risk. Neither did the court note the actual studies pointing to the disproportionate targeting of minorities. It is in the direction of those studies, and those countering them, that our analysis now turns.

Studies that claim RAs are racially biased

In 2016, staff writers for Pro-Publica, a non-profit investigative journalism organization dedicated to "shining a light on abuses of power and betrayals of public trust"[29] published "Machine Bias," an analysis of RA tools.[30] With the teaser line, "There's software used across the country to predict future criminals. And it's biased against Blacks," Pro-Publica fueled the debate about the role of RA in perpetuating harsher treatment of African Americans in the criminal justice system. The authors of the article analyzed 7,214 cases of arrestees from Broward County, Florida who had been assessed with the COMPAS tool. They followed this group for two years and reviewed official arrest data to determine the relationship between an arrestee's COMPAS score and future offending. Their analysis focused only on people assessed with COMPAS at the pre-trial stage. They reported that overall, COMPAS correctly identified recidivists 61% of the time and violent recidivists 20% of the time. Black and White offenders were equally likely to

receive correct predictions of recidivism, but Black offenders had a higher rate of actual recidivism. They conclude that COMPAS has low predictive validity or is only "slightly better than chance" at predicting future criminality. Beyond the lack of accuracy in prediction, they claim the difference in prediction errors between Black and White offenders in their analysis supports a claim of racial bias. While the rate of prediction errors was similar for Black and White offenders, the types of errors differed significantly and disadvantage Black offenders. That is, COMPAS performs equally well in identifying Black and White offenders who recidivate, but when it errs, it overestimates the number of Black offenders who will recidivate and underestimates the number of White offenders who will do so. They report that Black offenders were incorrectly classified as high risk (false positives) almost twice as often as White offenders (44.9% Black v. 23.5% White). Conversely, White offenders were incorrectly classified as low risk (false negatives) more often than Black offenders (28% Black v. 47.7% White). In other words, they argue their analysis indicates that inaccurate COMPAS predictions are more likely to unjustly result in erroneous predictions of high risk and increased penalties for African Americans, who are not a threat to public safety, and leniency for Whites who do, in fact, pose a threat. Like the Wisconsin Supreme Court, they note that Northpointe, now Equivant, creators of COMPAS, would not disclose the calculations used to create COMPAS scores on proprietary grounds, thus shielding the tool from analysis and critique.

Studies that do not identify racial bias in RAs

The ProPublica article was published online with color photos and compelling anecdotes of individuals suffering injustices because of judicial reliance on COMPAS scores. It was reported in the *New York Times*, *The Washington Post*, the *Atlantic* and a PBS story on

the uses and misuses of algorithms. In response, several scientists reviewed the data and analysis made available by ProPublica on their website. Northpointe also responded to the article, pointing out flaws in calculations and data presentations.[31] The Northpointe reanalysis of the data used by ProPublica came to the opposite conclusion; White offenders were more likely to be incorrectly classified as high risk than Black offenders (41% White v 37% Black) and Black offenders were more likely to be incorrectly classified as low risk (29% Black v 35% White). They identify inappropriate classification statistics as the source of the incorrect attribution of racial bias.

Flores and colleagues challenged ProPublica's conclusions on methodological and analytical grounds.[32] Briefly, they suggest ProPublica's sample, approach to classification, interpretation of mean score differences, disregard for accepted standards of tests for bias, and inflated tests of significance raise concerns about the validity of the analysis.[33] Flores et al. re-analyzed ProPublica's data and reached the opposite conclusions. They found COMPAS accurately predicted recidivism for White and Black offenders at the same rates, and found no evidence of racial bias in predictions for this particular sample of pre-trial defendants in Broward County, Florida.[34]

Some critics have argued that risk markers that reflect selective bias mask racial discrimination that occurs earlier in a defendant's life.[35] For example, one's criminal history may be influenced by racially biased decisions. African Americans are disproportionately represented at all stages of the criminal justice system, so any selective bias that exists in arrest, charging and sentencing will be reflected in criminal history. Some have claimed that RAs are basically race assessments and that bias is essentially "baked in" to criminal history.[36]

Former Attorney General Eric Holder noted the distinction between static and dynamic factors when he raised concerns about RA in his 2014 speech to the National Association of Criminal Defense Attorneys. While noting the benefits of RA, particularly for men of color, he also warned of their potential to "undermine our efforts to ensure individualized and equal justice".[37] He discouraged consideration of static factors over which one has no control, such as current education level, socioeconomic background or neighborhood, and rather emphasized the importance of the characteristics and circumstances of the current charge and a defendant's prior criminal history. Other analysts have suggested that static factors beyond a defendant's control or clearly correlated with race or ethnicity, such as poverty or family structure, are not appropriate for inclusion in RA and contribute to the potential for racial and gender bias.[38] Some suggest that inclusion of dynamic risk factors, such as anti-social attitudes and peers who support criminal conduct, are important for devising effective interventions, especially with youth.[39]

Most IPV RAs rely heavily on static factors, such as prior threats to kill and use of weapons, due to their focus on the history of abuse. However, Connor-Smith et al. (2011) found that inclusion of dynamic factors, such as escalating violence, combined with victims' assessments, increase usefulness of RAs both for prediction and victim safety planning. Since some forms of IPV, namely intimate terrorism[40] or coercive control,[41] typically comprise an ongoing, course of conduct crime, it is important that IPV RAs appreciate the changing nature of risk, potentially assessing it through screenings at multiple junctures/episodes.

Jennifer Skeem and Christopher Lowenkamp provide further clarification on the issue of racial bias in RAs.[42] They analyzed the risk scores, recidivism and race of 34,794 federal offenders assessed with the Post Conviction RA (PCRA). Like most research in this area, they compared Black and White offenders and did not examine other racial or ethnic categories, arguing that racial disparities in the criminal justice system primarily affect African Americans. Relying on standards developed for other domains, specifically the *Standards for Educational and Psychological Testing*, they found no evidence of racial test bias in the PCRA for predicting recidivism. They explain that mean score differences between groups do not equate with bias. African American defendants, as a group, score higher on the PCRA than White defendants principally due to more extensive criminal histories among African Americans. Two-thirds of the differences in mean scores between Black and White defendants is attributable to prior criminal history. Criminal history predicts future arrests of Black and White offenders with equal accuracy.[43] Further, by examining the contributions to risk scores of race and criminal history, together and separately, they determined that criminal history mediates the relationship between race and risk scores. Race, on its own, is a very weak predictor of PCRA scores, until it is combined with criminal history. A history of arrest for violent crimes is the strongest factor contributing to high PCRA scores. Criminal history does not "stand in" for or serve as a proxy for race.[44] The focus on selective biases in criminal history is not consistent with this analysis and overlooks the reality that criminal history is already included in the sentencing guidelines of most jurisdictions, separate from its consideration in a RA.[45]

While mean score differences do not equate with test bias, they may still contribute to disparate impact in sanctions. Higher mean scores can contribute to higher rates of

incarceration for Black offenders. It is vital that data analyses follow standard scientific procedures, but questions about the fairness of RA instruments cannot be answered through mathematics alone. Human beings must weigh various factors in decisions about fair practices. These decisions involve trade-offs among accuracy and fairness, defendant rights and victim safety.

RA, ethics, and fairness

A number of scholars have identified the challenges of adhering to standards of fairness in employing RA tools. First, individuals enter the criminal justice system from a variety of social locations that carry historical and contemporary advantages and disadvantages. Economically, racially and ethnically marginalized populations are more likely to enter the criminal justice system than privileged groups and to receive harsher sanctions. Although poverty and racism do not cause crime, they generate conditions conducive to criminal activity.[46] Rates of offending and recidivism vary and are higher among low-income and African American people. People disagree about the sources of this variation, with some pointing to consistency between self reports, victim reports and official data[47] and others emphasizing biased selection processes involving disproportionate targeting of groups, including intensive policing, implicit racial biases, and explicit discrimination against African Americans.[48] Whatever the source, these varying offending/recidivism rates impact the risk scores produced by any instrument that includes prior criminal history, the strongest predictor of future offending. Indeed, some have argued that youth and prior criminal history are such powerful predictors that addition of other risk markers adds very little to predictive accuracy. Policy makers who select tools based on their ability to predict recidivism will choose those that include prior

criminal history and thus will choose accuracy over minimization of racial and economic bias.

Fairness in RAs is not a simple matter. There are different types of fairness, and they are not all compatible. A number of authors have explained the impossibility of achieving all types of fairness simultaneously in a context where two groups, Black and White offenders, have different rates of known recidivism.[49] Three types of fairness in RA focus on whether all subgroups within a population of people assessed have: calibration or predictive parity; equal rates of false negatives; and equal rates of false positives. Calibration means that risk scores would have equal rates of correct predictions across people from all groups with that risk score. This was the case for the results of the COMPAS study conducted by ProPublica. Among White offenders classified as high risk, 60% reoffended; among Black offenders classified as high risk, 61% reoffended. This "predictive parity," or equal accuracy in predictions between groups, means that the COMPAS tool met the calibration criteria of fairness. While not a particularly impressive success rate, this level of success in predicting reoffending is considered legitimate in the field of risk prediction. When applied to pre-trial decisions, it will result in thousands of people who score as low risk being released on their own recognizance rather than remaining in jail. Virginia and New Jersey have adopted pre-trial RA tools and have experienced a significant reduction in jail detainees with no accompanying increase in officially measured crime. However, it is not possible to maintain predictive parity and achieve equal rates of false positives and false negatives between groups that differ in their rates of reoffending. Because Black offenders have higher rates of prior arrests, and thus more Black offenders have high-risk scores, they are at greater risk of being incorrectly identified as high risk.

That is, Black offenders are more likely to be in the category of false positives. Conversely, since White offenders are more likely to be classified as low risk, they are more likely to be incorrectly identified as low risk, that is, to be in the category of false negatives. The only mechanisms for eliminating these second two types of unfairness are to lower the threshold at which White offenders would be identified as high risk, or raise the threshold at which Black offenders are identified as high risk. Either option would violate equity notions of fairness and also result in more dangerous offenders being released simply because they are Black.[50] Another possibility would be to eliminate criminal history from the calculation of risk, which would significantly decrease the predictive accuracy of the tool. If Black and White offenders had identical rates of reoffending, RA tools would be able to achieve all three types of fairness. Since this is not the case, ethical decisions must be made about which types of fairness to sacrifice.

The relative cost of false positives and false negatives is one consideration in making these decisions. The obvious cost of false positives is potential increases in surveillance and control of people who are apparently not a threat to public safety.[51] This can harm people's livelihoods, relationships, education, health and self-esteem and result in harsher sentences than those given to people who are not detained prior to trial. If the cost of false positives falls more heavily on one particular group, for example, African Americans, it undermines people's faith in the justice system and contributes to general cynicism about the social order and the rule of law. On the other hand, the cost of false negatives is release and minimal supervision of offenders who will commit future crimes. If these offenses are limited to non-violent and drug offenses, the primary cost does not entail bodily harm. If they are violent offenses, the cost is physical harm, potentially homicide. This includes IPV

offenders incorrectly classified as low risk and released with little or no supervision and who pose a serious and ongoing threat to their partners and possibly other community members.

We should not consider the inherent incompatibility of these forms of fairness as insurmountable barriers to the socially beneficial, ethical use of RA tools. Policy makers should examine any RA tool for racial bias. Some jurisdictions now require "Racial Impact Statements" that assess the racial and ethnic impact of new policies.[52] There are alternatives to incarceration, even for high-risk offenders, which can contribute to the goal of reducing jail and prison populations while protecting the public. When dynamic risk factors, such as employment and substance abuse, are included in RAs, they provide directions for treatment that have proven successful. Coupled with heightened supervision, intervention into the needs of offenders can reduce recidivism. Diversion programs that include electronic monitoring and frequent assessments demonstrate strong potential for helping people refrain from criminal activity.

Critics of RA tools have focused on the general offender population. Here we are concerned with the potential for bias in IPV RAs. Victims of IPV are in a different relationship to the person who harms them than are victims of other non-relational forms of crime. Those involved in determinations of fairness in assessing IPV offenders must be aware of the dynamics of IPV and the power differentials between victims and offenders. In the late 1970s and 1980s, the battered women's movement struggled to obtain legal protection for victims of IPV. Civil suits, such as Thurman v. City of Torrington, CT,[53] federal commissions and consultations, such as the US Commission on Civil Rights, *Battered Women: Issues of Public Policy*,[54] and academic research, such as the study of police in

Chicago[55] all revealed the lack of protection afforded victims of IPV. Marjory Fields of Brooklyn Legal Services opined: "As a class, battered women are denied the protection afforded to other victims of crime. They are discriminated against by police, prosecutors, and judges. As women victims of crime, they are not believed."[56] Over the past forty years, advocates, activists, policy makers, criminal justice practitioners and scholars have transformed the criminal justice response to battered women. Through sustained efforts to document the unique experiences of victims of IPV, we now have laws and policies that recognize the vulnerability of victims and the danger posed by abusive partners. When we consider fairness to perpetrators in this context, we must ensure that the rights afforded to this particular class of crime victims are not compromised.

IPV RAs and the complex question of racial bias

The development of IPV RAs

IPV RA tools emerged out of 50 years of varying types of research and a growing recognition of the complexity of the relationships evidencing IPV and particularly its more dangerous manifestations in intimate terrorism and coercive control. IPV RAs resemble other risk assessments insofar as they seek to incorporate actuarial or statistical data into the predictive process. Such incorporation increases the likelihood of accurately predicting outcomes. Actuarial approaches add an element of objectivity, consistency, and reliability across cases and case types, improving upon impressionistic assessments conducted by clinicians such as psychologists, counselors, police officers, or victim advocates.[57]

The danger of abstraction from bigger issues

Nevertheless, the science is far from foolproof, positive predictive capabilities relatively low, independent evaluations and validations of the effectiveness of these tools

almost non-existent, and, for our specific purposes, questions of racial bias essentially unaddressed. Indeed, insofar as IPV RAs function to triage cases, highlight potential danger, and recommend specific interventions in individual cases, they can be seen as contributing to making sense of IPV at the level of individuals and case types, rather than as a much broader social structural, political, moral, and cultural problem. Put differently, the language of risk abstracts danger from political arenas. This abstraction is particularly dangerous when we are talking about communities of color, particularly African-American and American Indian communities, where contemporary concentrated poverty looms large, historical disadvantage weighs heavy, and where IPV and IPH rates are substantially higher. Indeed, as anthropologist Mary Douglas comments, "Instead of isolating risk as a technical problem we should formulate it so as to include, however crudely, its moral and political implications."[58] Importantly, she opines, "The experts on risk do not want to talk politics lest they become defiled with political dirt," and "Indeed, reading the texts on risk it is often hard to believe that any political issues are involved."[59]

None of these caveats mean these tools do not have great potential for triaging cases, assisting victims, holding offenders accountable, and improving coordinated community responses through the development of a shared language of risk and danger. However, we must bear in mind that the issue of racial bias is not merely a technical question, one to be resolved through various statistical proofs or by considering whether a victim/alleged victim, once receiving the IPV RA wisdom, goes down this choice chamber or that. Rather our questions about the rational choices of actors (e.g. IPV victims, the accused, individual judges, victim advocates, and so on) ought be augmented with critical questions about the array of institutions, agencies, stakeholders, and state provisions

charged with addressing IPH and severe IPV, phenomena that are socially patterned not merely a product of individual behavior and choices. Indeed, just as good risk assessors emphasize the importance of assessing the meaning and contexts within which specific risk profiles emerge in individual cases, so too is it essential to ask about the context and meaning of IPV RAs themselves.

Risk and gender

Historical studies consistently report men commit a greater share of IPH than women in the U.S, although the male-to-female ratios of perpetration vary by race and/or ethnicity and over time.[60] The historical record notes an even greater overrepresentation of male perpetrators in homicide-suicides and familicides than in cases of single decedent IPHs.[61] These findings about the gendered nature of IPH comport with those of Dobash and colleagues who remark, "Men often kill wives after lengthy periods of prolonged physical violence accompanied by other forms of abuse and coercion; the roles in such cases are seldom if ever reversed."[62] Most tools therefore predict risk to female victims, a focus that reflects the research that generated the tool in the first place, the fact that female victims are more likely to experience outcomes such as severe re-assault, intimate terrorism, coercive control, and IPH, and the political milieu within which IPV RAs arose and persist.

The potential idiosyncrasies of IPH

As research studies proliferated, we learned more about what happened in IPHs. To date, the focus has been almost exclusively on men killing current/former spouses or intimate partners. Descriptive statistics mapped the frequencies of what loosely came to be called "risk factors" in IPHs.[63] Increasingly, researchers noted the frequencies of prior histories of domestic violence, an actual or pending separation or emotional estrangement

between the intimate partners, violent and controlling jealousy on the part of the man, and so on. Reassuringly, these frequencies corresponded to the findings of the growing network of domestic violence fatality review teams across the US and elsewhere.[64] Nevertheless, at the turn of the century, one researcher cautioned, "The research shows little if any qualitative difference in the antecedents to lethal and non-lethal domestic violence."[65]

Soon, pioneering comparative studies sought to distinguish between cases of IPH and cases where men abused women but did not kill them. In Campbell et al.'s landmark retrospective (not prospective) study, the research team compared 220 cases of IPH with cases of 343 abused women that formed a comparison or referent group. For Campbell et al., "A woman was considered 'abused' if she had been physically assaulted or threatened with a weapon by a current or former intimate partner during the past 2 years."[66] Different patterns emerged between the death and control group. For example, murdered women were more likely than controls to have been assaulted or threatened with a weapon, had partners who had access to guns, been subject to strangulation and forced sex, subject to violence that was increasing in frequency and/or severity, beaten during pregnancy, and threatened with death. These disproportionate correlations of "risk markers" with the IPHs signaled the possibility that these were some of the hallmarks of more dangerous IPV cases that might be used for triaging.[67]

Some potential pros and cons of IPV RAs

Other important studies tracked the outcomes of IPV RAs over time, providing insights into the combination and permutation of questions that might possibly predict future severe re-assault in various settings. Messing and colleagues used a quasi-experimental field trial and concluded that administering the Lethality Assessment

Protocol (LAP; an IPV RA) was associated with an increase in victims taking protective actions (e.g. removing or hiding a partner's weapon, sought formal domestic violence services) and a decrease in the frequency and severity of violence among their sample of IPV survivors.[68]

IPV RA tools help victims understand the potential danger they face. If it turns out to be true that the use of IPV RAs reduces re-victimization and enhances self-protective behaviors, they may indeed prove very useful. However, we simply cannot predict human behavior very easily.[69] Of those positive screens for future severe re-assault, a relatively low percentage of victims actually experience the adverse outcome (e.g. severe re-assault) within the specified time period. This may be (guardedly) good news for victims who test positive. However, the implications for the accused, if the tool is somehow used against them in the criminal justice system, are potentially profound and raise serious due process issues. A couple of examples from current risk tools help flesh out these concerns.

Technical matters with implications for due process: relative and absolute risk and positive predictive values

Snider et al.'s research found that a five question IPV RA (Danger Assessment-V) might work well in hospital emergency departments.[70] They deemed that affirmative answers to three out of five questions meant that a previous IPV victim was at "high-risk" of future severe re-assault within 9 months of her baseline interview. However, of those high-risk victims answering three of the five questions affirmatively, only 25 percent actually experienced severe re-assault within that time frame.[71] Risk assessors refer to the probability that those testing positive will actually experience the adverse outcome within a predefined time period as the "positive predictive value (PPV)." Low PPVs characterize

IPV RAs in general. To date, these low predictive values have not received due attention. We will return to the significance of these relatively low PPVs.

The recently developed Arizona intimate Partner Risk Assessment Instrument System (APRAIS) is another case in point with a low PPV.[72] A victim that answers four or more out of seven RA questions affirmatively has a 10.5 times greater chance than someone answering zero or one risk questions "yes" of experiencing severe re-assault within seven months. Four or more "yes" responses place the victim/alleged victim (choose your language) in the "high-risk" APRAIS category. As a triaging mechanism with large numbers of calls, police and victim advocates find the APRAIS potentially useful. Victims report finding it a helpful way of assessing the potential dangers they face.

Nevertheless, the absolute risks victims face are very different. Among victims who answer four or more APRAIS predictive questions "yes", only roughly 15 percent will experience severe re-assault within the ensuing seven months. One can see, depending on the context within which the APRAIS is used, possible due process concerns looming large.

Will the long term field operationalization of our two example IPV RA tools, both of which purport to measure the probability of future severe re-assault, (the Danger Assessment-5 (PPV of 25% at the 3/5 "yes" threshold for high-risk) and the APRAIS (PPV of 15% at the 4/7 "yes" threshold for high-risk), result in the disproportionate identification of the accused who are African American? [73] Unlike the COMPAS tool, researchers have no data on these matters, although we obviously need that data.

Pre-trial settings

Currently, there is no validated IPV RA for use in pre-trial settings. As noted, the IPV RAs screen for recidivism, not whether the accused might return for a hearing or pose a

threat to the community in the intervening period (i.e., for pre-trial outcomes). Put simply, the research in this arena has not really commenced. As a corollary, the possibility of detecting racial bias in the use of IPV RAs pre-trial awaits research.

Notwithstanding our observations about the nascent state of the art regarding IPV RAs, we also note, for the record and from our field experience, that the problem of understanding the impact of IPV RAs at pre-trial is more daunting than it may first appear. The recent utilization of the APRAIS in Arizona provides a case in point. At time of writing, roughly 25 of Arizona's 174 law enforcement agencies have, to varying degrees, begun deploying the APRAIS at domestic violence crime scenes. With the consent of victims/alleged victims, officers administer the APRAIS once their investigation is complete and the question of probable cause resolved. Officers notify victims of the discoverability of the risk information they might disclose. Assessments are then (ideally) attached to the law enforcement release questionnaire (charging sheet) and presented to the court at the initial appearance of the accused. So far, we have no data regarding whether the rate at which victims decline to participate in the RA differs by race, or, what any such differential participation rates might mean.

Hypothetically, in some Black communities with concentrated poverty, victims/alleged victims of IPV might be less likely than Caucasian peers in wealthier neighborhoods to participate in a RA, regardless of whether Black victims/alleged victims are more or less likely to actually call police in the first place. This reluctance may stem from their fear of being seen to cooperate with law enforcement. Such reluctance may result in Black victims not being connected with victim advocates and support services. Such disparate receipt of services may be correlated with the overall deployment of the IPV

RA in a community but the reasons for such disproportionate service utilization may not be directly related to the deployment of the tool itself but rather the broader historical and socially situated realities of its field operationalization.

Tracking the relationship between, for example, high-risk APRAIS scores and bail determinations, release of offenders on their own recognizance, future recidivism, future police calls to the residence for domestic violence, and so on might appear straightforward. For example, it might appear straightforward to examine the relationship between pre-trial decision-making (e.g., bail levels, release on the accused's own recognizance), APRAIS scores and race of the accused and/or victim. Indeed, tracking correlations between these variables is possible. However, drilling down on why a particular judge set a certain bail amount in cases where the judge received an APRAIS addendum is inordinately complex. Our fieldwork involves asking judges how they make bail determinations in IPV cases. In Arizona judges have up to 15 factors to consider, one of which may include an APRAIS score. When asked, judges are not necessarily able to articulate what weighting they attach to an APRAIS high-risk score versus other factors. The topic requires considerable research, but it is impossible to escape the seeming reality that many judicial decisions regarding pre-trial release are complex, intuitive, constrained by the realities of tight dockets, and therefore highly subjective. Arizona judges are not required to disclose how they weigh each of the up to 15 factors. It would therefore be very challenging to determine the impact of APRAIS scores on bail setting, let alone tease out whether the APRAIS scores had a differential impact by the race of the accused or the victim/alleged victim, or both.

Other settings and contexts: trial, sentencing, and the civil arena

It is also possible for prosecutors to use IPV RAs at stages beyond pre-trial, e.g., trial and sentencing. Here the possibility of racial bias also warrants careful consideration. Even if we had anything more than anecdotal data on these matters from individual prosecutors, it lies beyond our remit to examine these important possibilities.

We must also consider the possible use of IPV RA tools in the civil justice arena, for example, in regard to divorce, order of protection hearings, and child custody. Research suggests that such tools may be very useful in providing judges with more comprehensive information about IPV risk markers, above and beyond what appears in petitions/affidavits for relief through orders of protection.[74] It is essential to explore whether in these arenas, the use of IPV RAs and their specific risk markers/questions results in racial bias. Clearly, judges in the civil arena, for example, family court judges, are required to use state law as their principal frame of reference. IPV RAs do not constitute evidence and should therefore not be used either to grant or deny orders of protection. But as recent case law suggests, risk markers, although not the type of facts properly taken by judicial notice, may still be used as a partial frame of reference for issuing orders of protection.[75]

The dangers of neglecting political dirt

Even if it was the case that IPV RAs identify the precise levels of danger posed by an abuser, it is important to remember that they do not and cannot identify the principal sources of danger within the life of an individual victim of IPV. Threats to victims take many forms, including acute poverty, historic and biographical traumas of sometimes inestimable import, lack of health care, affordable housing, and childcare, diminished life chances in an increasingly deregulated and globalized market place, restricted access to reproductive care, limited advocacy and community support services, and intimate

partners subject to a growing array of disadvantages that at times likely feed a rage, shame, and deeply compromised masculinity not well addressed in batterer intervention programs. Herein lies what many may see as the real dangers to battered women. This is not to diminish the threats posed by individual abusers but rather to resituate them amidst what Douglas calls "political dirt."

Assume IPV RAs were able to provide us with PPVs of 100%, precise delineations of the threat posed by batterers. So what? How can we best utilize such information? What can we do about victim danger in those relationships? In jurisdictions across the US, IPV RAs have revealed that between 45% and 80% of the assessments put victims at "high" or "elevated" risk.[76] Early data returns from the deployment of the APRAIS in one county in Arizona reveal almost 60 percent of victims to be at high or elevated risk. Many of those Arizona victims have been connected with local domestic violence service providers, utterly overwhelming those providers, who in good faith and due diligence have been paying to house them in hotels as stopgap measures.

Regardless of the predictive capabilities of the APRAIS tool, the highly personal (yet legally discoverable) risk testimonies mean the high risk victims (nearly all women) have experienced four or more of the following: violence that increases in frequency and/or severity; ongoing, violent, and extreme sexual jealousy; an abuser the respondent thinks is capable of killing them; an abuser who has beaten the victim while the victim was pregnant; a perpetrator who has used a gun, object, or other weapon against the victim; a perpetrator who has previously tried to kill the victim; a perpetrator who has strangled, choked, or suffocated the victim. The question is not whether we can assess violent outcomes in the lives of these victims with any great degree of technical accuracy or

mathematical precision. Neither is it whether the IPV RA tools are racially biased, although this is undesirable, immoral, and so on. Rather, a more important question to ask might be: Why do Black and American Indian women suffer alarmingly high rates of intimate terrorism and IPH? Why is roughly one quarter of the Phoenix, Arizona Sojourner Center shelter residents African-American when only 2 percent of Phoenix's population is Black?

There is a dangerous tendency to blame a victim once that victim receives information about the risk of future violence she faces at the hands of her abuser. Whether IPV RA tools are racially biased matters much less than the racial realities of diminished life chances that plague minorities like African Americans and American Indians. Useful as they may be at the level of individual victims and cases, IPV RAs can never compensate for the historical and social traumas they cast their eyes over. Indeed, once risk is assessed and the victim duly notified, blame may wait in the wings, albeit implicitly, couched in the language of personal choices and options, amidst shrinking resources and harsher times.

Conclusion

In summary, we note that IPV RAs differ from RAs in general. The research into the outcomes of using the former is in its infancy. We simply do not have the data to answer the question of whether IPV RAs are racially biased. Whatever position researchers adopt on whether RAs in general are racially biased, researchers, jurists, practitioners, and others appear of similar mind in recognizing the undesirability of racially biased assessment tools and procedures. From our field practice we have provided a glimpse into just how difficult it might prove to determine whether IPV RAs are racially biased. Such research would prove expensive, time consuming, and involve an intricate blend of quantitative and qualitative methodologies and ways of knowing about case dynamics and case flow. To the

extent IPV RAs communicate danger through a shared language, inculcate wariness among victims, agencies, and stakeholders, and provide rational frames of reference in non-biased ways, they have tremendous promise at the level of individual cases. However, if IPV RAs enable an ever-more frugal deployment of supports and interventions within a risk society and a punitive and increasingly polarized culture of control, they may become part of the problem not part of the solution.

---

[1] Schmidt, Caitlin, "Man accused in two slayings to undergo competency exam due to erratic courtroom behavior." Arizona Daily Star, July 18, 2018. https://tucson.com/news/local/man-accused-in-slayings-to-undergo-competency-exam-due-to/article_43ded337-0e2f-59d4-b87c-c149e3c63532.html. Accessed September 5, 2018.
[2] Khmara, Danyelle, "Death Sentence." *Tucson Weekly*, Jan. 18, 2018. https://www.tucsonweekly.com/tucson/death-sentence/Content?oid=14066948. Accessed September 5, 2018.
[3] See Beck, Ulrich, 1992, *Risk Society: Towards a New Modernity*, London, Sage; Beck, Ulrich, 2016, *The Metamorphosis of the World*, London, Polity.
[4] See Garland, David, 2001, *The Culture of Control*, University of Chicago Press, Chicago, Illinois.
[5] See Violence Policy Center reports for long term declines in IPH; see Bureau of Justice Statistics, November 2012, Intimate Partner Violence, 1993-2010, page 1, Figure 1 for details of decline in overall IPV from 1993.
[6] Garland, 2001, p. 11.
[7] Garland, op. cit., p. 11.
[8] Garland, op. cit., p. 11.
[9] Harcourt, Bernard E., 2008, *Against Prediction : Profiling, Policing, and Punishing in an Actuarial Age*. Chicago: University of Chicago Press.
[10]Hannah-Moffat, K., & Maurutto, P., 2010, Re-contextualizing pre-sentence reports. *Punishment & Society,* 12(3), 262-286. doi:10.1177/1462474510369442; Flores, A.W., Bechtel, K. & Lowenkamp, C.T., 2016, *Federal Probation* 80(2), 1-36; Kleinberg, J., Lakkaraju, H., Leskovec, J., Ludwig, J., & Mullainathan, S., 2018, Human decisions and machine predictions. *Quarterly Journal of Economics, 133*(1), 237-293. doi:10.1093/qje/qjx032; Skeem, J. L., & Lowenkamp, C. T. , 2016, Risk, race, and recidivism: Predictive bias and disparate impact. *Criminology, 54*(4), 680-712. doi:10.1111/1745-9125.12123.

[11] Hannah-Moffat, K., 2012, Actuarial sentencing: An "unsettled" proposition. *Justice Quarterly* 30(2), 270-296.
[12] Andrews, D.A., Bonta, J., Wormith, J. S., 2011, The risk-needs-responsibility (RNR) model: Does adding the good lives model contribute to effective crime prevention? *Criminal Justice and Behavior* 38(7), 735-755.
[13] Mayson, S. G., 2018, Dangerous defendants. *Yale Law Journal,* 127(3), 490-568; Silver, E., & Miller, L., 2002, A cautionary note on the use of actuarial risk assessment tools for social control. *Crime & Delinquency, 48*(1), 138-161. doi:10.1177/0011128702048001006
[14] Tonry, M., 2014, Legal and ethical issues in the prediction of recidivism. *Federal Sentencing Reporter*, 26(3), 167. doi:10.1525/fsr.2014.26.3.167

[15] Angwin, J., Larson, J., Mattu, S., Kirchner, L., Machine Bias, ProPublica, 2016, May 23. 2016, https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing, accessed September 5, 2018 ; Starr, S. B., 2014, Evidence-based sentencing and the scientific rationalization of discrimination, *Stanford Law Review, 66*(4), 842-872; Harcourt, B. E. (2008). *Against prediction: Profiling, policing, and punishing in an actuarial age*. Chicago: University of Chicago Press; Holder, E. 2014, Speech presented at the National Association of Criminal Defense Lawyers 57th Annual Meeting and 13th State Criminal Justice Network Conference, Philadelphia, PA., 2014, *Federal Sentencing Reporter* 27(4), 252-255.

[16] See Wiseman, S., 2009, "Discrimination, Coercion, and the Bail Reform Act of 1984: The Loss of the Core Constitutional Protections of the Excessive Bail Clause, *Fordham Urban Law Journal* 36, 1: 121 - 157.

[17] Browder, K. , 2018, Bail reform and risk assessment: The cautionary tale of federal sentencing. *Harvard Law Review, 131*(4), 1125-1146.

[18] Jurisdictions vary on the time limits within which an initial appearance must occur but it cannot be excessive.

[19] State of Wisconsin v. Loomis, 881 N.W.2d 749 (WI 2016).

[20] The US Supreme Court declined to hear the case on June 26, 2017. According to one journalistic interpretation prior to the declination, "If it hears the case (Loomis), the court will have the opportunity to rule on whether it violates due process to sentence someone based on a risk assessment instrument whose workings are protected as a trade secret," Rebecca Wexler, "When a Computer Program Keeps You in Jail," *New York Times*, June 13, 2017. However, the brief filed by the State of Wisconsin urged the US Supreme Court to deny Loomis's petition in part because the "use of risk assessments by sentencing courts is a novel issue, which needs time for further percolation" (Brief in Opposition, page 6). Unlike Wexler's interpretation, attorneys for the state of Wisconsin agreed that the trial court "considered" the COMPAS assessment but did not base the sentence upon that assessment.

[21] *Loomis*, 881 N.W.2d at 754.

[22] Correctional Offender Management Profiling for Alternative Sanctions (COMPAS).

[23] *Loomis*, 881 N.W.2d at 756.

[24] *Loomis*, 881 N.W.2d at 757.

[25] *Harvard Law Review (HLR)*, Volume 130 at 1532, Criminal Law – Sentencing Guidelines – Wisconsin Supreme Court Requires Warning Before Use of Algorithmic Risk Assessments in Sentencing.- *State v. Loomis*, 881 N.W.2d (Wis. 2016).

[26] *Loomis*, 881 N.W.2d at 769.

[27] *Loomis*, 881 N.W.2d at 769.

[28] HLR, 1533.

[29] ProPublica webpage, https://www.propublica.org/about/. Accessed September 11, 2018.

[30] Angwin, op. cit.

[31] Dieterich, W., Mendoza, C., Brennan, T., 2016, COMPAS Risk Scales: Demonstrating accuracy equity and predictive parity. July 8. Northpointe. http://go.volarisgroup.com/rs/430-MBX-989/images/ProPublica_Commentary_Final_070. Accessed September 11, 2018.

[32] Flores, Bechtel and Lowenkamp, op. cit.

[33] ibid., p. 9-10

[34] ibid., p. 22

[35] Harcourt , op. cit.; Silver and Miller, op. cit.; Starr, op. cit.

[36] Harcourt , op. cit.; Tonry, op. cit.

[37] Holder, op. cit., p. 254.

[38] Starr op cit.; Tonry op. cit.

[39] Peterson-Badali, M., Skilling, T., & Haqanee, Z., 2015, Examining implementation of risk assessment in case management for youth in the justice system. *Criminal Justice and Behavior, 42*(3), 304-320. doi:10.1177/0093854814549595.

[40] See Johnson, Michael, 2008, *A Typology of Domestic Violence*, Boston, MA, Northeastern University Press.

[41] See Stark, Evan, 2007, *Coercive Control: How men entrap women in personal life*, New York, Oxford University Press.
[42] Skeem and Lowenkamp, op. cit.
[43] Skeem, J. and Monahan, J., 2011, Current directions in violence risk assessment. *Current Directions in Psychological Science,* 20(1): 38-42, p. 701.
[44] Ibid., p. 700.
[45] Frase, R. S., 2014, Recurring policy issues of guidelines (and non-guidelines) sentencing; risk assessments, criminal history enhancements, and the enforcement of release conditions. *Federal Sentencing Reporter, 26*(3), 145-157. doi:10.1525/fsr.2014.26.3.145
[46] Walker, S., Spohn, C., and DeLone, M., 2011, *The Color of Justice: Race, Ethnicity and Crime in America,* 5th ed., Belmont, CA: Wadsworth, p. 99, quoted in Skeem and Lowenkamp, op. cit., p. 685.
[47] Skeem and Lowenkamp , op. cit.
[48] Harcourt, op. cit.; Starr, op. cit.
[49] Berk, R., Heidari, H., Jabbari, S., Kearns, M., & Roth, A., 2017, Fairness in criminal justice risk assessments: The state of the art. Preprint retrieved from http://arxiv.org/abs/1703.09207, September 10, 2018; Kleinberg, J., Mullainathan, S., & Raghavan, M., 2016, Inherent trade-offs in the fair determination of risk scores, Proceedings of Innovations in Theoretical Computer Science, Preprint retrieved from http://arxiv:1609.05807 , September 10, 2018; Chouldechova, A., 2017, Fair prediction with disparate impact: A study of bias in recidivism prediction instruments. *Big Data, 5*(2), 153-163. doi:10.1089/big.2016.0047; Corbett-Davies, S., Pierson, E., Feller, A., Goel, S., & Huq, A. (Aug 4, 2017). Algorithmic decision making and the cost of fairness. Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge, Discovery and Data Mining, pp. 797-806. doi:10.1145/3097983.3098095 Retrieved from http://dl.acm.org/citation.cfm?id=3098095, September 11, 2018.
[50] see Petersilia, J., & Turner, S., 1987, Guideline-based justice: Prediction and racial minorities. *Crime and Justice, 9*, 151-181.
[51] Silver and Miller, op. cit.
[52] Skeem and Lowenkamp , op. cit.; Mauer, M., Racial impact statements as a way of reducing unwarranted sentencing disparities, 2007, *Ohio State Journal of Criminal Law*, 5(1), 19-46.
[53] 595 F.Supp. 1521 (C.D.1 1984)
[54] United States Commission on Civil Rights, 1978, *Battered Women: Issues of Public Policy, A Consultation Sponsored by the United States Commission on Civil Rights.* January 30-31. Washington, D.C.: U.S. Commission on Civil Rights.
[55] Parnas, R., 1967, The police response to the domestic disturbance, *Wisconsin Law Review,* 914, 914-960.
[56] US Commission on Civil Rights, op. cit., p. 21.
[57] Actuarial IPV RAs are much better at predicting future violence than clinical or professional judgments. See Hilton, N.Z., et al, 2008, An in-depth actuarial assessment for wife assault recidivism, *Law and Human Behavior*, 32, 150-163.
[58] Douglas, Mary, 1992, *Risk and Blame: Essays in Cultural Theory*, Routledge, London, page 51.
[59] Douglas, 1992, 38.
[60] For a general discussion of these ratios see Dobash R., Emerson and Dobash, Russell P., 2015, *When Men Murder Women*. Oxford University Press; Websdale, N., *Understanding Domestic Homicide*, Northeastern University Press, Boston, MA, 1999; and Wilson, Margo I., and Martin Daly, 1988, *Homicide*. Aldine de Gruyter, New York, 1988.
[61] See for example Websdale, N., *Familicidal Hearts*, Oxford University Press, London, 2010; Cohen, D., 1995, "Homicidal compulsion and the conditions of freedom," *Journal of Social History*, 28, 4: 725-764.
[62] Dobash, Russell P., R. Emerson Dobash, Margo Wilson, Martin Daly,1992, "The Myth of Sexual Symmetry in Marital Violence," *Social Problems*, 39 (1): 71-91, p. 81.
[63] Websdale, 1999, op. cit.
[64] See Websdale, N., Celaya, A. & Mayer, S., 2017, The development of domestic violence fatality review in the U.S., pp. 27-58 in *Domestic Homicides and Death Reviews: An International Perspective*, edited by Myrna Dawson, Palgrave, MacMillan, New York.

[65] Websdale, N., 2000, "Lethality Assessment Tools: A critical analysis," VAWNET, Violence Against Women Grants Office Applied Research Series.
[66] Campbell, J., et al., 2003, "Risk factors for femicide in abusive relationships: Results from a multisite case control study," *American Journal of Public Health*, 93, 7: 1089-1097, p.1089.
[67] It lies beyond our remit to explore the various criticisms leveled at Campbell et al.'s research. Suffice it to say that legitimate concerns exist that the comparative study over-selects the IPHs because IPH is rare and the control or comparison group contains a highly disproportionately low number of cases. Concerns might also be raised about the use of proxy informants to speak on behalf of murdered women. The researchers used a low number of proxy informants, perhaps limiting their ability to comprehensively learn about some characteristics such as whether the deceased victim thought her abuser was capable of killing her.
[68] Messing, J. T., Webster, D., Brown, S., Wilson, J.S., 2015 "The Oklahoma lethality assessment study: A quasi-experimental evaluation of the Lethality Assessment Program," *Social Service Review*, 89(3): 499-530.
[69] In *Barefoot v. Estelle* 463 U.S. 880 (1983) the US Supreme Court ruled that expert testimony on dangerousness may not always be correct but that it should be admissible and subject to the adversarial process. The court did not disagree with the amicus brief of the American Psychiatric Association that pointed out that in two thirds of cases, psychiatric predictions of future dangerousness were wrong. For our purposes, this translates into a PPV of 33%. *Barefoot* concerned sentencing, specifically, whether to invoke the death penalty. Put simply, the stakes were high. IPV RAs are used at earlier stages in the criminal justice system, e.g., by police officers at crime scenes, by victim advocates, and, in some states, e.g., Maine and Arizona, by judges at the initial appearance of the accused.
[70] Snider, C., Webster, D., O'Sullivan, C., & Campbell, J., 2009, "Intimate Partner Violence: Development of a Brief Risk Assessment for the Emergency Department," *Academic Emergency Medicine*, 16: 1208-1216.
[71] See Table 6, page 1213.
[72] Websdale, N., et al., 2018, Protecting Victims of Intimate Partner Violence: Arizona's Emerging Risk Assessment Model, *Police Chief*, April, pages 42-47.
[73] Snider, C., Webster, D., O'Sullivan, C., & Campbell, J., 2009, op. cit.; Websdale, N., et al., 2018, op. cit.
[74] Nichols-Hadeed, C., et al., 2012, Assessing Danger: What Judges Need to Know, Family Court Review, 50, 1, January 2012, 150-158.
[75] Pettinghill v. Pettingill, 480 S.W.3rd 920 Kentucky Supreme Court, 2015.
[76] Messing, J., Campbell, J., Wilson, J.S., Brown, S. & Patchell, B., 2017, The Lethality Screen: The Predictive Validity of an Intimate Partner Violence Risk Assessment for Use by First Responders, *Journal of Interpersonal Violence,* 32(2), 205-226; Kretschmer, J., Bukach, A., Bray, R., 2018, Cuyahoga County Domestic Violence High Risk Team One Year DA-LE Report. March. Begun Center for Violence Prevention Research and Education, Case Western Reserve University, Cleveland, Ohio.