Secret algorithms that predict future criminals get a thumbs up from Wisconsin Supreme Court

(Ethan Chiel, splinternews.com, July 27, 2016)
There’s software used across the country to predict future criminals. And it’s biased against blacks.

Julia Angwin, Jeff Larson, Surya Mattu and Lauren Kirchner, ProPublica, May 23, 2016.

Risk Assessment – A Brief Look Back

• We’ve all been engaged in risk assessment since the dawn of time
• Humans, and practically all other organisms, are constantly scanning our environment for potential risks
• We come up with schemas, heuristics, and shortcuts – informal algorithms - that allow us to more quickly identify and then either avoid or manage risks accordingly
  • For example, we likely learned as children to look twice both ways before we cross a road
Risk Assessment – A Brief Look Back, cont.

• Application of algorithms to risk assessment is not something new – it has however become more formalized and widespread due to advances in computing power and acquisition of vast quantities of information – Big Data
Risk Assessment – What is it?

For our purposes today, we will define risk assessment as estimating the likelihood that an offender will recidivate (revert to illegal behavior) after the individual has experienced legal consequences and/or an intervention for a prior criminal act.
Purposes of Risk Assessment

- Risk assessment spans the entire spectrum of the adjudication process
  - Case Processing:
    - Arrest and Investigation
    - Adjudication
    - Sentencing and dispositional purposes (e.g., pretrial confinement, sentencing, custody level of institutional placement or supervision level of community placement)
Purposes of Risk Assessment, cont.

- **Corrections:**
  - Determination of intervention needs, settings, and programming

- **Re-entry and Monitoring:**
  - Supervision
  - Community-based treatment and other interventions
  - Sex Offender Registration and Notification (SORN) and civil commitment proceedings

- **Community Support**
  - Prevention and education
  - Victim services and outreach
  - Community involvement
The Importance of Risk Assessment

The effectiveness of efforts to manage and treat offenders is predicated in many respects on the accuracy of risk assessment — our ability to accurately differentiate offenders based on their assessed risk for recidivism.
Risk Assessment – A Brief History

- Hanson’s Developmental Stages of Risk Assessment (Hanson, 1998)
  - Unguided (or unstructured) clinical judgment
  - Guided (or structured) clinical judgment
  - Research-guided clinical judgment
  - Pure actuarial approach
  - Adjusted actuarial approach
Risk Assessment – A Brief History. cont.

• Bonta’s (1996) generational model describing the development of risk assessment approaches:
  • **First Generation**: Unstructured professional opinion
  • **Second Generation**: Actuarial methods using static predictors
  • **Third Generation**: Methods that incorporate both static and dynamic predictors of risk
  • **Fourth Generation**: Assessment of risk and case management planning to address and ameliorate risk
Comparison of Risk Assessment Approaches

Comparisons of the above approaches have a long and at times contentious history. While the superiority of structured to unstructured approaches appears to have been settled (Grove, 2005; Grove et al., 2000; Meehl, 1954), each of the structured approaches has its merits and both supporters and detractors.
Types of Risk Factors

• Static Risk Factors
  • Immutable, relatively unchangeable factors
    • If they do change, they typically only increase
    • Generally not amenable to intervention efforts

• Dynamic Risk Factors (aka Criminogenic Needs, Psychologically Meaningful/Relevant factors)
  • Stable
  • Acute
  • These are often amenable to intervention
Accuracy of Risk Assessment

- The accuracy of risk estimates depends in part on two critical factors:
  - The degree to which the individual offender being assessed matches a known group of offenders, and
  - The degree to which the factors included in the risk assessment accurately reflect the known universe of relevant risk factors
Accuracy of Risk Assessment, cont.

• With actuarial risk assessments, the utility of a specific risk factor depends on its empirical relationship to the outcome being predicted.
• Base rates must be taken into account; we arrive at base rates generally through large meta-analyses of offender recidivism.
• No single best predictor of risk has yet been identified.
Common factors associated with risk have been identified however, and appear across a wide range of risk assessments. A well-known list of these factors, compiled though decades of criminological research, is the Central Eight risk factors identified by Andrews and Bonta (2006)
The Central Eight (Andrews and Bonta)

1. History of antisocial behavior
2. Antisocial personality
3. Antisocial cognitions
4. Antisocial associates
5. Substance abuse
6. Poor familial/marital relationships
7. Poor performance at work and/or school
8. Lack of prosocial involvement and activities
Consideration of Factors

• The accurate assessment of risk involves gaining an understanding of as many available, relevant factors associated with the known criterion or outcome behavior as is possible and/or practical.

• While important to consider dynamic factors when assessing risk, doing so via an adjustment of actuarial instruments may not be the most effective way.
Consideration of Factors, cont.

• Studies that have examined the effects of making clinical adjustment to actuarial risk assessments have found that overrides - a consideration of factors outside the actuarial scheme - decrease predictive accuracy (Hanson, 2009; Hanson & Morton-Bourgon, 2009).

• These studies (Gore, 2007; Hanson, 2007; Vrana, Sroga, & Guzzo, 2008) have all been prospective in nature, and they involved actuarial instruments currently used with sex offenders.
How Many Measures Should I Use?

Two primary rationales support the notion that using more than one instrument provides potential benefits:

- Classical test theory suggests that increasing the number of items in an assessment increases reliability and coverage.
- If there are multiple driving forces behind offending behavior, and individual risk assessment instruments tap these underlying dimensions or pathways to offense recidivism differentially, then the use of multiple instruments would have a distinct advantage over the use of a single instrument alone.
Communication of Risk Assessment Results

Descriptive categories: Individuals have widely varying perceptions of what is meant by the terms we use to describe risk. For instance, one person’s high may be another’s medium risk. It also depends on the consumer’s frame of reference – an offender may be high risk compared to non-offending members of the public (the consumer’s natural point of reference), but low risk relative to similarly situated offenders (the appropriate frame of reference). These terms also have different meanings from one measure to the next.
From Hanson et al. (2017) “Researchers compared risk-level definitions among five assessment measures and found that only 3 percent of the people assessed were identified as high risk across all five instruments and only 4 percent of the people were identified as low risk by all five measures. This means that the same person can be described by different categories across different assessment instruments, or people in the same category can be described differently across different assessment instruments.”
Communication, cont.

- Numerical estimates may not be readily understood by consumers of risk assessment information.
- Concepts such as odds ratios or percentiles may not be as familiar to the consumers of risk information as they are to those who develop and administer risk assessments.
Hanson’s efforts in this arena (Hanson et al., 2017) to arrive at a common language around communication of risk assessment are supported by BJA.

They recommend a five-level scheme of communicating risk, with five descriptive categories that are each associated with a specific set of characteristics and risk of reoffending.
Current Controversies in Risk Assessment

- As noted in our introduction, risk assessment has been in the news a lot recently, and not always for the best reasons. Generally, these concerns or controversies fall into a few distinct categories, which we will explore in turn:
  - The **development** of risk assessments
  - The **use** of risk assessments
  - **Constitutional issues** regarding risk assessments
  - **Algorithms and Big Data**, and their role in risk assessments
Development of Risk Assessments

• The data we use – do the data exacerbate disparities?
• The factors we choose – do they violate constitutional clauses? Are we introducing demographic and socioeconomic factors that would otherwise be prohibited from consideration due to constitutional concerns?
• The degree to which risk assessments are a “black box” – opaque, versus transparent. What are the implications?
Use of Risk Assessments

• At what point in the adjudication process are we conducting a risk assessment, and for what purposes?
  • Pretrial risk assessment is relatively straightforward, in that there are two fairly obvious outcomes of interest – the likelihood of Failure to Appear and the risk to public safety, defined as behavior resulting in an arrest while on pretrial release.
  • While risk is the central issue in pretrial, in sentencing it is only one of the many considerations to be made, and may be neither central nor determinative in nature.
Use of Risk Assessments, cont.

Risk assessments in sentencing are more controversial, given that sentencing involves two crucial decisions – the type of sentence and the length of sentence. These decisions are informed by different theories of punishment that address concepts such as individual retribution, rehabilitation, deterrence, and incapacitation. Application of risk assessment beyond pretrial is fraught with potential complications, and should be undertaken with a great deal of thought and care.
Use of Risk Assessments, cont.

• Evidence-Based Sentencing (EBS) — the latest proposed revisions to the Model Penal Code endorses the use of risk assessments in the shift to Evidence-Based Sentencing

• As of 2017, five states require the use of risk assessments in sentencing, albeit in different ways (Arizona, Oklahoma, Kentucky, Ohio, and Pennsylvania)

• Ohio and Pennsylvania seem to have taken a more thorough approach (significant consultation and guidance from experts and academics) to the use of risk assessment in sentencing relative to the other states, who have in many ways left it up to the judiciary to sort out the issues
Use of Risk Assessments, cont.

The Justice Department under the previous administration, specifically former AG Eric Holder, urged that we take a cautious approach to the use of algorithms in risk assessments used for sentencing purposes:

“Experience and analysis of current risk assessment tools demonstrate that utilizing such tools for determining prison sentences to be served will have a disparate and adverse impact on offenders from poor communities already struggling with many social ills.”
Use of Risk Assessments, cont.

Holder: “Although these measures were crafted with the best of intentions, I am concerned that they may inadvertently undermine our efforts to ensure individualized and equal justice. By basing sentencing decisions on static factors and immutable characteristics – like the defendant’s education level, socioeconomic background, or neighborhood – they may exacerbate unwarranted and unjust disparities that are already far too common in our criminal justice system and in our society.”
Use of Risk Assessments, cont.

- Holder urged that we use extreme caution when sentencing criminals using predictive analytics: “on the possibility of a future crime that has not taken place” - rather than on the specific facts of the crimes.

- Days prior to Holder’s speech, DOJ’s Criminal Division sent a letter to the Chair of the U.S. Sentencing Commission expressing concerns about including predictive analysis in criminal sentencing decisions, noting that risk assessment instruments “raise constitutional questions because of the use of group-based characteristics and suspect classifications in the analytics.”
Constitutional Issues

- The use of risk assessment algorithms raises two primary concerns from a constitutional perspective (both 14th Amendment issues):
  - The individual’s right to due process – in the context of risk assessment, this involves having access to the information that was used to arrive at a decision or judgment so that you can in essence confront your accusers as well as the issue of applying group norms and associated recidivism probabilities to an individual defendant.
Constitutional Issues, cont.

*Equal protection clause* – It has been argued by Sonia Starr of the University of Michigan Law School that including risk assessments in criminal sentencing is tantamount to “an explicit embrace of otherwise condemned discrimination, sanitized by scientific language.” Her reasoning is that by permitting consideration of risk assessments, we are allowing judges to consider factors such as age, gender, and socioeconomic status that have long been considered inappropriate to consider at sentencing.
Constitutional Issues, cont.

• Although the Equal Protection Clause does not require that the government treat every person exactly the same, it does prohibit discrimination if it is based upon impermissible classifications.

• Washington v. Davis (426 U.S. 229; 1976) found that while not irrelevant, disproportionate impact is permissible as long as the law or practice was adopted *absent the intent* of racial discrimination. Courts have found however that it is not permissible to base sentences on gender.
Algorithms and Big Data

Algorithm has recently become a dirty word. Just a few years ago, it seemed that as a society we were for the most part enamored of algorithms and enthusiastic about the promise of Big Data. Lately however concerns about the degree to which we have allowed algorithms to “control our lives” has become a concern – the fear of algorithms appears to have become merged somewhat with the fear of Artificial Intelligence – that we will be overtaken by smart machines, a la 2001: A Space Odyssey or I, Robot.
Algorithms and Big Data, cont.

• These fears seem due at least in part to some of the revelations from Edward Snowden and Wikileaks, as well as the increasing incursions of Google, Amazon, and Facebook into more and more areas of our lives – all of which rely heavily on algorithms and Big Data

• We seem to readily accept the role played by algorithms in other areas of our lives, such as in the healthcare and insurance industries
Algorithms and Big Data, cont.

• The concerns regarding algorithms and Big Data generally concern the use of risk assessment algorithms in sentencing, rather than other applications. Specific issues include:
  • Opacity – the lack of transparency: The risk here is that tools developed by for profit companies are proprietary and lacking in transparency for trade purposes. By contrast, those developed by academics and researchers in conjunction with government agencies are less of a black box – the factors and weights are generally made available, and they are therefore more open to scrutiny.
Algorithms and Big Data, cont.

• Bias and lack of reliability/validity: The 2016 ProPublica study of the COMPAS, which suggested that the COMPAS was both inaccurate in its risk assessments as well as racially-biased, despite not explicitly including race in its algorithm (The publisher of the measure, Northpointe (d/b/a equivant), has vigorously disputed the findings described in the ProPublica study).
Algorithms and Big Data, cont.

A growing number of legal scholars and academics are concerned that the use of algorithmic risk assessments may not only reinforce, but may actually exacerbate existing racial and ethnic disparities within the criminal justice system.
Algorithms and Big Data, cont.

Cathy O’Neil, in her book *Weapons of Math Destruction*, argues that bias is inherent in our data, such as in arrest records that reflect both “reported” and “found” crimes. Policing patterns impact these two types of offenses differently, in that found crimes are more likely to occur in areas where police observation is more frequent – such as poor and minority neighborhoods.
Algorithms and Big Data, cont.

- It is critical to analyze the different types of error found, and to decide what types and how much error you are willing to accept
  - These choices are to be made in consultation with policy and decision makers, and are context-dependent
- The types of error have to be examined and monitored to ensure that they are free of bias – this is one of the central issues addressed in the ProPublica study of the COMPAS
• Concepts of Fairness: Legal fairness prioritizes **equality or parity in the process** by which an outcome is reached, rather than fairness of the **outcome itself**.
Where Do We Go From Here?

Rather than revert back to clinical risk assessment (remember the findings of Grove and Meehl; Bonta and Andrews; and Hanson and his collaborators), I advocate for the responsible development and use of actuarial risk assessment procedures where possible. Increasingly, these will move from paper-and-pencil measures to automated, actuarial algorithms that rely, at least in part, on Big Data.
Where Do We Go From Here, cont.

Objective, research-based tools can help reduce the biases that result from the large amounts of discretion historically applied throughout the criminal justice system.

We know a great deal more about criminal behavior and its antecedents than we did fifty years ago, and have access to much more data and computing power.
Where Do We Go From Here, cont.

We can’t pretend that bias (both inadvertent and overt) has not for decades been a part of criminal justice processing and judicial decision making. That’s no excuse however to institutionalize such bias through the use of risk assessment algorithms. Therefore we need to exercise great caution moving forward.
Consider that the same factor that might increase assessed risk may also serve a mitigating function. In *Penry v. Lynaugh* (492 U.S. 302; 1989), the US Supreme Court found that the defendant’s intellectual limitations simultaneously increased risk and reduced his blameworthiness for the crime committed.
Given the amount of attention recently concerning the issue of potential bias in the risk assessment process, as well as our own experiences with the development and validation of risk assessments, I offer some suggested guidelines to ensure that we as a field are intent on minimizing potential sources of bias and respecting constitutional issues in our work developing risk assessments.
Guidelines to Consider

1. Be open and transparent regarding the included factors and how the factors are chosen and weighted
   • Facilitate regular outside research and auditing of the tools and associated processes
   • Be open to considering other factors and weights, based on input from relevant stakeholders

2. Ensure that the risk assessment is being used correctly, for its intended purpose, and applied to appropriate populations – for instance, normed using local populations on which the measure is used
3. Address error types and associated rates with policy and decision makers, and ensure that the rates do not differ based on demographic factors

4. Continually test risk assessments to ensure that they perform equally well across all groups to which you are applying the measure
   • If the measure either over- or under-predicts for a certain sub-population, that suggests bias
   • Risk assessments need to demonstrate equivalent levels of reliability and validity (especially predictive validity) across all subgroups
   • If a measure does not, re-formulate and re-validate the measure until it does or create separate measures for each sub-group
Guidelines to Consider, cont.

5. Re-validate the measure as conditions dictate (and approximately every five years regardless)

6. Stay current on training and developments in the risk assessment field

7. Understand that gender and class are linked with so many other issues
Guidelines to Consider, cont.

8. From Kehl et al., (2017) – the concept of “Technological Due Process” – ensuring that ample opportunity exists to challenge the results of risk assessments
   • Core values are: transparency, accuracy, accountability, participation, and fairness
   • The presence of a clear “audit trail” – involving the ability to track decisions, open access to source data and the actual algorithms
9. Policy makers need to be involved in and aware of all steps in the risk assessment process.

10. Demonstrate a strong commitment to fairness
Guidelines to Consider, cont.

11. To the extent possible, use the most parsimonious set of factors possible, eliminating controversial factors and their proxies whenever possible

• O’Neil argues that we should exclude troubling variables even if doing so decreases the predictive accuracy of our algorithms

• Ensure that strong procedural safeguards are in place to ensure as much as possible that the scores are used properly and that their inadvertent impact is minimized
Guidelines to Consider, cont.

12. Communicate risk assessment results effectively, consistently, and in a manner that to the greatest extent possible leads to the responsible use of risk assessment information.
The Future of Risk Assessment

The genie is out of the bottle – despite reasonable, significant, and meaningful concerns, the use of algorithms in risk assessment is only likely to increase. As a result, as developers, practitioners, and policy makers, we need to ensure that their use is ethical and constitutionally sound. A commitment to the above concepts of transparency, accountability, and fairness will be critical.
The Future of Risk Assessment, cont.

Risk assessment is not now, nor will it ever be perfect, unless we somehow come to a point where we are able to collect all the information on all the relevant factors, and know precisely how to combine them optimally to provide a perfect level of prediction. We must therefore accept a certain amount of error as part of the process.
Demanding perfection in our methods is a higher bar than what we accept in practically every other aspect of our lives. For instance, open heart surgery results in death on the operating table in about 1 - 3% of cases. That’s certainly a less than perfect outcome. Should we then abandon open heart surgery because it isn’t perfect?
The Future of Risk Assessment, cont.

In the end, we should be committed to using the best information and science we have, choosing the most appropriate and accurate assessments, applied in a responsible manner that does not extend them beyond their appropriate and ethical use, communicate the results in as effective and standardized a means as possible, while always keeping in mind the limitations and fallibility of the process and working on ways to improve our tools and methods.
Questions and Discussion
Kevin Baldwin, Ph.D.
Applied Research Services, Inc.
3235 Cains Hill Pl NW
Atlanta, GA 30305
kbaldwin@ars-corp.com
www.ars-corp.com
Questions?

Thank you for attending today’s webinar. All of JRSA’s public webinars are available on our webinar webpage located at

http://www.jrsa.org/webinars/index.html
All telephones have been muted. If you would like to ask a question please use the chat feature unless instructed otherwise. Please remember to select **Host, Presenter or Panelists**.

During the webinar you may also be asked to select **All Participants**. This will allow you to share your answers with your fellow participants.
Help Us Count

If you are viewing as a group, please go to the chat window and type in the name of the person registered and the total number of additional people in the room (e.g. Michelle Smith, 5).

This will help us track the number of people attending our webinars.