Research-Based Risk Assessment:
Adding Equity to CPS Decision Making

By:
Christopher Baird
Janice Ereth, PhD
Dennis Wagner, PhD

June 1999

Children's Research Center

426 Yellowstone Drive, Suite 250   voice (608) 831-1180
Madison, WI 53719   fax (608) 831-6446

Table of Contents

A.   INTRODUCTION

Over the last two decades, America's Child Protective Service (CPS) systems have seen

the number of African American families on caseloads increase. In most jurisdictions, African

Americans are more likely than Whites, Hispanics, or Asians to have allegations of maltreatment

substantiated and to have children placed in out-of-home care. In many states, the difference

between these rates for Whites and African Americans is substantial.

This trend is particularly disturbing when viewed in the context of the National Incidence

Studies (NIS) which conclude that there are no differences in maltreatment rates for African

Americans and Whites.[1] If these studies are accurate, they raise serious questions about the

disproportionate number of African American families in America's child protection system.

Some have worried (or even concluded) that as more and more CPS agencies implement

actuarial risk assessment systems, racial bias will be exacerbated. They reason that because

actuarial systems use simple objective factors such as income levels, family size, and number of

caretakers in the home, more African Americans than Whites will be rated high risk. Because

more African American families than Whites live in poverty, have more single caretaker

households, and have three or more children residing in the household, these fears, at least on the

surface, seem to have some foundation. If bias exists, it could result in additional over

representation of African Americans in both the CPS and foster care systems.

However, rather than speculate about the potential impact of research-based systems on

the over representation of African Americans, the Children's Research Center (CRC) has

---

[1]The National Incidence Studies develop estimates of abuse and neglect prevalence rates for the general population. These studies combine data from several sources, including child protection services and professionals from non-CPS agencies likely to come in contact with maltreated children. Participants included law enforcement, schools, day care centers, hospitals, voluntary social service agencies, mental health agencies, juvenile probation, and public health departments.

reviewed data from agencies using actuarial risk assessments to determine if these systems have resulted in a greater level of disparity between African Americans and Whites. The number of agencies that have implemented actuarial systems has increased considerably in recent years. As a result, risk ratings on many thousands of families, both African American and White, are available for analysis. This paper presents data from the three largest states using CRC actuarial risk assessment systems, California, Georgia, and Michigan. Data are also presented from a national study of risk assessments where the Michigan risk assessment system was applied to cases from four geographically dispersed jurisdictions. These data presentations are followed by an in-depth discussion of results, as well as issues raised regarding risk assessment and the disproportionate representation of African Americans in child welfare.

This paper will answer questions about the relationships between risk, race, and recurrence of abuse and neglect. It presents the actual experiences of states using actuarial risk assessment and clarifies precisely how equity issues should be evaluated.

B.   RACE AND RISK INFORMATION FROM STATES USING ACTUARIAL RISK ASSESSMENTS

Alaska was the first state to use actuarial risk assessment in CPS and has now done so for more than a decade. Soon after Alaska implemented its system, Michigan followed suit and has since accumulated a massive amount of data on family risk levels, needs, services, and outcomes. From the beginning, both developers of the risk assessment system, CRC and Michigan officials, were highly cognizant of potential racial bias and made equity a key principle of the development effort. Risk instruments were tested independently on families of each race to ensure that equity was maintained. The results are clear; despite using factors that some have worried are "biased,"

essentially equal proportions of African Americans and Whites are classified to each risk level in Michigan.

Table 1 presents risk assessment results from assessments of nearly 12,000 Michigan cases. Obviously, the degree of equity attained in classification of Michigan families to risk levels is quite remarkable.

Table 1		
Michigan **Percent of Families at Each Risk Level**		
Risk Level	**Whites (N=6,651)**	**African Americans (N=5,296)**
Low	10.5%	11.3%
Moderate	30.7%	30.0%
High	45.1%	46.0%
Very High	13.7%	12.7%

Another simple method for exploring the equity issue is to compute an average risk level for each group.[2] Using this approach to compare risk ratings of different groups again demonstrates that the Michigan risk assessment tools do not discriminate against African Americans. African Americans had an average risk rating of 2.601, while Whites averaged 2.620. In sum, Whites, not African Americans, had a marginally higher risk rating when the actuarial assessment was applied.

The Michigan experience is no anomaly. In fact, findings from all states using actuarial risk assessments are notably consistent. Georgia adopted the Michigan system, modified it to a

---

[2]This is done by assigning a value to each risk level, multiplying the value by the number of cases at each level, summing the totals and dividing by the total population of each group. The following values were assigned to Michigan risk levels:

    Low Risk = 1; Moderate Risk = 2;   High Risk = 3; Very High Risk = 4

small degree, and then implemented it statewide. Data are now available on more than 4,000

assessments conducted in a sample of Georgia counties. Table 2 outlines the results of these

assessments by race. Again, more Whites than African Americans score at the higher risk levels.

Using "average risk ratings" for comparison purposes, Whites in Georgia have an average risk

rating of 1.852, while the average rating for African Americans is 1.763.

Table 2		
Georgia **Percent of Families at Each Risk Level**		
Risk Level	**Whites (N=2,170)**	**African Americans (N=1,949)**
Low	27.5%	31.6%
Moderate	60.2%	59.5%
High	12.4%	8.9%

California began implementation of a Structured Decision Making (SDM) system this year

in several pilot counties. Actuarial risk assessment is a key component of this system. California,

with assistance from CRC, developed its own risk assessment instrument. Development was

based on a random sample of more than 2,500 families from seven pilot counties.[3] The research

resulted in risk assessment tools similar in both content and format to the Michigan system.

Classification results by race, using data from the scale construction effort, again illustrate that

actuarial risk assessment models do not systematically rate African Americans to higher risk levels

(see Table 3). Using the "average risk rating" of each race to compare results indicates again that

Whites, not African Americans, have a slightly higher risk profile. Whites in California averaged

---

[3]The counties included as pilots were Alameda, Humboldt, Los Angeles, Orange, Sacramento, San Bernardino, and Santa Clara.

2.582 while African Americans had an average rating of 2.337. The California data base also provides one of the best observations available to date on how risk levels of Hispanic families correspond to those of Whites and African Americans. As Table 3 illustrates, risk ratings of Hispanic families do not exceed those of Whites. The average risk rating for Hispanic families in the California study was 2.343.

	Table 3 **California** **Percent of Families at Each Risk Level**		
Risk Level	**African American** **(N=413)**	**Hispanic** **(N=682)**	**White** **(N=1,103)**
Low	17.4%	13.8%	11.5%
Moderate	39.0%	47.2%	39.2%
High	32.7%	29.9%	37.3%
Very High	10.9%	9.1%	12.1%

In a national study funded by the Office on Child Abuse and Neglect (OCAN) comparing different risk assessment systems, the Michigan system was applied to cases from Florida, Missouri, California, and Michigan. These data serve as the final example. As Table 4 illustrates, there were no significant differences in the number of Whites and African Americans scoring at each risk level and again, the average risk rating computed for each group indicates that Whites have a slightly higher overall risk profile (2.147 for Whites vs. 2.105 for African Americans).

Table 4		
Michigan Risk Assessment Applied to Cases from Four Jurisdictions Percent of Families at Each Risk Level		
Risk Level	Whites (N=684)	African Americans (N=581)
Low	13.5%	15.2%
Moderate	58.3%	58.8%
High	28.2%	25.9%

In sum, jurisdictions using actuarial risk assessments do not assign more African American families to higher risk levels. This, however, by itself is not sufficient to guard against bias. African American and White families rated high, moderate, and low risk must also exhibit similar rates of subsequent maltreatment. If, for example, a system rated 25% of both races "high risk" but 50% of high risk Whites had a subsequent report of abuse or neglect reported while only 30% of high risk African American families again abused or neglected children, then the term "high risk" means something very different for the two races. While minor differences in rates of subsequent abuse/neglect reports and substantiation are unavoidable, the instruments should 1) demonstrate a significant level of discrimination between high, moderate, and low risk cases within each racial group, and 2) never create a situation where high risk cases of one race do not exhibit significantly higher rates of subsequent abuse or neglect than moderate risk cases of any racial group. Moderate risk cases must, in turn, exhibit significantly higher rates of subsequent abuse/neglect than rates found for low risk cases of any racial group. When CRC joins with a state to develop a CPS risk assessment system, **it is careful to ensure that differences between races in maltreatment rates recorded at each risk level are minimized.** Not only do these systems classify similar proportions of cases of each racial/ethnic

group assigned to each risk level, each group exhibits similar rates of subsequent maltreatment

within risk levels. Data from the California and Michigan scale construction samples are

presented in Table 5.[4]

Table 5				
Subsequent Substantiation Rates by Race and Risk Level				
Risk Level	**Michigan****		**California****	
	African Americans	**Whites**	**African Americans**	**Whites**
Low/Moderate*	6.0%	5.0%	9.0%	15.0%
High	15.0%	12.0%	29.0%	33.0%
Very High	28.0%	30.0%	42.0%	50.0%

*   Because of the small number of cases rated low risk (when the sample is subdivided by race), the low and
      moderate categories have been combined.
**   Michigan rates reflect a 12-month follow-up. California rates reflect a 24-month follow-up.

---

[4]Georgia adopted the Michigan scale; since no retrospective scale construction research was conducted, no
outcome information is yet available.

C. THE PURPOSE OF RISK CLASSIFICATION SYSTEMS

Before discussing why actuarial risk instruments increase equity in decision making, it is important that the CPS field has a full understanding of the purpose of these systems. The field should move beyond traditional notions of what risk assessment is all about. Validity of decision systems has until lately been measured by the degree to which "predictions" about case outcomes are realized. For example, Ruscio (1998) defines validity in the following manner:

> "The efficacy of your decision policy can be examined through the use of a simple fourfold classification table crossing the optimal outcome for each child (kept at home vs. placed into care) with the decision that is reached. There are two types of correct decisions, or "hits," that are possible: True positives are decisions that place children into care when appropriate, and true negatives are decisions that keep children at home when appropriate. There are also two types of incorrect decisions, or "misses," that are possible: False positives are decisions that unnecessarily place children into care, and false negatives are decisions that fail to place children into care when placement is necessary. Based on this classification table, the effectiveness of a decision policy may be evaluated in several ways. For instance, one could determine how many of the decisions to place a child into foster care were correct (true positives divided by the sum of true and false positives); how many children who optimally should have been kept in the home actually were (true negatives divided by the sum of true negatives and false positives); or how many placement decisions, overall, were correct (the sum of true positives and true negatives divided by the total number of cases)."

While calculations of false positives, false negatives, and the overall percentage of correct predictions is useful in many settings, it may not be the best method for gauging the efficacy of a risk assessment system when the probability of success/failure is substantially different than 50-50. When events are relatively rare, they are inherently difficult to predict. In such instances, simply assuming an event will not occur may produce more predictive accuracy than any attempt to determine where or when occurrence is likely. For example, if subsequent maltreatment (failure) is reported in only 15% of cases opened to protective services, then simply predicting <u>no case</u> opened to services will have subsequent maltreatment reported produces an 85% "hit rate."

Obviously, such a prediction, while highly accurate, is of little value to a CPS agency. (The "sensitivity" of the prediction is .85, but the specificity -- correct identification of those who do fail -- is zero.) A valid and reliable risk assessment system may improve the "hit rate" marginally, but it is possible such a system could result in a <u>higher</u> percentage of false positives and false negatives and still provide the agency with quality information about the relative probability of subsequent maltreatment. Consider the scenario where a CPS population (N=100) has a subsequent maltreatment rate of 15%. A risk assessment identifies 25% of the population as "high risk," which, for this example, is equated with a prediction of subsequent maltreatment. Actual versus predicted outcomes are presented below:

<table>
<tr><td colspan="4" align="center">**Table 6**

Example
Actual vs. Predicted Outcomes</td></tr>
<tr><td></td><td></td><td colspan="2" align="center">**Predicted Outcomes**</td></tr>
<tr><td></td><td></td><td align="center">**No Subsequent Maltreatment**</td><td align="center">**Subsequent Maltreatment**</td></tr>
<tr><td rowspan="2" align="center">**Actual Outcomes**</td><td>**No Subsequent Maltreatment**</td><td align="center">71</td><td align="center">14</td></tr>
<tr><td>**Subsequent Maltreatment**</td><td align="center">4</td><td align="center">11</td></tr>
</table>

In the above example, an overall "hit rate" of 82% is attained (3% lower than that attained when all cases are predicted to succeed) with a rate of false positives (subsequent maltreatment) of 56% and false negatives of 5.3%. Despite the high proportion of false positives, cases that were rated high risk experienced maltreatment at a 44% rate, while only 5.3% of those rated at lower risk levels had subsequent maltreatment reported. The ratio of "failures" in the high risk

group to "failures" in the low risk group is more than 8:1. Such results help agencies identify which families are more likely to abuse or neglect their children. In addition, 11 of the 15 cases (73.3%) where subsequent maltreatment occurred were correctly identified (a relatively high rate of specificity).

Other fields, such as medicine, juvenile justice, and adult corrections, have largely abandoned the idea that risk assessment is an exercise in prediction. Instead, terms like base expectancy rates have replaced discussions of false positives and false negatives. In corrections, for example, high risk does not equal a prediction of failure -- in fact, in most correctional systems, more high risk cases succeed than fail. Instead, high risk simply denotes inclusion in a group of offenders with significantly higher historical rates of recidivism than other groups.

The field of medicine offers similar examples. In cancer research, it is common practice to identify characteristics of malignancies and surrounding tissue and to classify patients as high, moderate, or low risk based on the observed rates of recurrence within a specified time period (see for example the Van Nuys Prognostic Index cited in Silverstein and Lagios, 1997). A designation of high risk of recurrence does not equate with a "prediction" that the cancer will recur. In fact, most medical professionals carefully avoid making such predictions. As treatment options expand and improve, recurrence-free survival rates have increased to the point where, if false positives and negatives were to be minimized, the best "prediction" for high risk cases would be "no recurrence." Still, knowing that cases with similar characteristics have experienced a recurrence rate of 10%, 25%, or 45%, helps the doctor and patient select the most appropriate

treatment plan. In medical literature, the use of stark terms such as false positives and false

negatives to evaluate the efficacy of cancer classification techniques is relatively rare.[5]

Conceptually, the use of false positives and false negatives to evaluate risk assessment

systems creates another dilemma. While outcomes are often dichotomous (an event will either

occur or not occur), most risk assessment systems assign cases to three or more different risk

levels. If system efficacy is based on predicting an outcome, it must be asked what prediction is

being made for cases at intermediate risk levels: Is the designation "moderate risk" a prediction

that subsequent maltreatment will or will not occur? We submit that it is neither, but simply the

recognition that these cases "recidivate" at higher rates than some and lower than others.

Knowing this allows workers to establish appropriate service plans, just as similar information

permits doctors and patients to decide on a particular course of action.

Therefore, in evaluating the relative efficacy of risk assessment instruments, it is

imperative to be very clear about expectations. The terms prediction and classification are often

used interchangeably, yet really connote different expectations. Prediction is more precise than

classification. According to Webster, prediction "declares in advance on the basis of observation,

experience, or scientific reason." To predict accurately in any field is difficult; to accurately

predict human behavior is especially complex as so many factors contribute to determining how

individuals will act. Classification, on the other hand, is simply "a systematic arrangement in

groups or categories according to established criteria." While accurate prediction would greatly

benefit CPS and society, it has not proven feasible. We submit that the goal of risk assessment is

---

[5] For example, in two recent research studies that produced or evaluated "scales" to rate the risk of recurrence of breast cancer (Silverstein and Lagios, 1996) (Silverstein and Lagios, 1997), the terms false positives and false negatives are not used. What is most interesting about these articles is that the scaling methods employed mirror those used to create many child welfare risk scales.

much more modest; it is simply meant to assign cases to different categories based on observed rates of behavior.

What is important is the <u>degree</u> to which families in different risk groups perform differently. Classification recognizes that a high risk designation is not a prediction of failure. It is, instead, a clear indication that such families may require more attention and more services, because cases in this designation tend to "fail" at higher rates than cases in other classifications. Valid risk instruments achieve significant differences in rates of maltreatment among risk groups. The greater the differences, the better the instrument.

Recently, better methods of measuring the efficacy of risk assessment systems have emerged. Researchers are realizing that traditional definitions of validity do not adequately measure the power of risk classification systems. Silver and Banks (1998), for example, state that "the primary utility of a risk classification model is in providing a continuum of risk estimates associated with a variety of conditions which can be used to guide a range of decision making responses . . . it is for this reason that traditional measures of 'predictive accuracy' which carry with them the assumption that dichotomous decisions will be made, have little utility for assessing the <u>potency</u> of a risk classification model."

In essence, the validity of risk assessment systems should be measured by the degree to which subgroups of a meaningful size are identified by each system and the degree to which different rates of subsequent maltreatment are reported for each subgroup.

D.   WHY ACTUARIAL SYSTEMS PRODUCE SUCH A HIGH DEGREE OF EQUITY

Data from states using actuarial risk assessment systems clearly demonstrate that these instruments do not discriminate against African American families. In fact, they bring a level of

equity to CPS decision making that could very well serve as at least a partial remedy to racial disproportionality that has plagued child welfare over the years.

Understanding why actuarial risk assessment systems work so well to increase equity in decision making requires a knowledge of how they are constructed, as well as a thorough understanding of national data on abuse and neglect as it relates to race. The following discussion links national data on abuse and neglect and methods used to construct systems for measuring risk. It is important that CPS agencies understand these relationships.

The National Incidence Studies (NIS), as noted earlier, indicate that there are no significant differences in maltreatment rates between African Americans and Whites. Therefore, it is reasonable to expect that there should be little difference in risk ratings for families of each race. As the data presented above demonstrate, for states using actuarial risk assessment systems, this is indeed the case. However, while the NIS have been prominently cited by those questioning the relationship between race and risk factors, NIS data really have little relevance to risk assessment research. The NIS combine several data sources to estimate prevalence rates **for the general population; abuse and neglect incidents are reported as occurrences per 1,000 people in a given location. Risk assessment studies, however, focus on only a small subset of the general population – those families who have had allegation of abuse or neglect reported and are then investigated. Recurrence rates** recorded for these families are much higher than incidence rates found in the general population (10 to 20 times the general incidence rate). The prevalence rates cited in the National Incidence Studies bear little resemblance to patterns of recurrence found in a CPS cohort. Despite this limitation, NIS data can be helpful in explaining why actuarial risk assessment systems do so much to promote equity in CPS decision making.

Although there may be no difference in the overall maltreatment rates reported for African Americans and Whites, a review of National Incidence Studies data base clearly indicates there are substantial differences in **type of maltreatment reported by race**. African Americans have a higher rate of neglect reported; Whites have a higher rate of abuse reported.

Maltreatment, particularly neglect, is strongly correlated with poverty, the stress involved in being a single parent, and other factors that are more commonly found in African American homes. For example, far more African Americans than Whites have annual incomes less than $15,000 (30.4% vs. 11.6%); a much greater percentage of African American homes have a single female caretaker present (52% vs. 18%); and a higher proportion of African American homes have three or more children present (14% vs. 9%).[6] The National Incidence Studies report higher rates of neglect for families with each of these characteristics:[7]

- Children in families with income less than $15,000 are 44 times more likely to be neglected than children from higher income families.

- Children in single caretaker families have an 87% greater risk of physical neglect than those in two caretaker families.

- Children in the largest families are neglected at three times the rate of single child families.

Given that families with these characteristics have a higher level of neglect reported and that more African American families have these characteristics, only one conclusion is possible: neglect rates are higher for African Americans than Whites. Still, it must be remembered that Whites facing the same stresses noted here also have similar incidence rates reported.

---

[6]U.S. Census Bureau (1995)

[7]NIS III (1996)

Conversely, the composition of White families makes the occurrence of child <u>abuse</u> more common than in African American families. There are several contributing factors, but the most prominent is gender related. As the National Incidence Studies report, abuse is most often perpetrated by males (67% of all physical abuse perpetrators are male; 89% of sexual abuse is perpetrated by males).[8] Less than half (48%) of all African American families have a male caretaker present. About 80% of White families have a male caretaker in the household.[9] With more families with a male caretaker present, Whites have higher rates of abuse reported.

The difference in rates of abuse and rates of neglect reported for African Americans and Whites essentially offset each other, **resulting in nearly identical rates of overall maltreatment**. However, the fact that these differences exist helps explain why actuarial risk assessment systems used in Michigan, Georgia, California, and other states effect such a high level of equity.

Most notably, in states using CRC risk assessment models, **separate instruments are used to rate the risk of neglect and the risk of abuse**. This is key to achieving equity as different family characteristics are related to the recurrence of different types of maltreatment. As data presented earlier would indicate, African Americans score higher on most neglect scale items and have a higher average risk of neglect score. Table 7 presents an item analysis from Michigan and from the four sites that participated in the OCAN study; similar results from other sites are also available. The differences in average neglect scores translate into higher proportions of African Americans rated high risk for <u>neglect</u> (see Figures 1 and 2).

---

[8]NIS III (1996)

[9]U.S. Census Bureau (1995)

	Mean Score			
Item	**Michigan**		**OCAN**	
	Whites	**African Americans**	**Whites**	**African Americans**
Current Complaint is for Neglect	.641	.743	.539	.617
Number of Prior Assigned Complaints	.912	.816	.916	.892
Number of Children in the Home	.373	.487	.349	.472
Number of Adults in Home at Time of Complaint	.359	.538	.356	.536
Age of Primary Caretaker	.317	.406	.340	.348
Characteristics of Primary Caretaker	1.21	1.23	.907	.812
Primary Caretaker Involved in Harmful Relationship	.546	.404	.470	.304
Primary Caretaker Has a Current Substance Abuse Problem	.589	1.064	.329	.613
Household is Experiencing Severe Financial Difficulty	.220	.267	.154	.151
Primary Caretaker's Motivation to Improve Parenting Skills	.727	.721	.403	.474
Caretaker(s) Response to Investigation	.803	.909	.553	.629
Average Neglect Score	**6.697**	**7.585**	**5.359**	**5.860**

Table 7

Neglect Item Scores from Michigan and the National OCAN Study

Figure 1

Neglect Risk Levels by Race
Michigan

25.3%

17.2%

7.1%

50.5%

27.0%

30.0%

3.0%

40.0%

African Americans

Whites

■ Low   ■ Moderate   □ High   ■ Very High

Figure 2

Neglect Risk Levels by Race
OCAN National Study

27.0%

59.0%

14.0%

33.0%

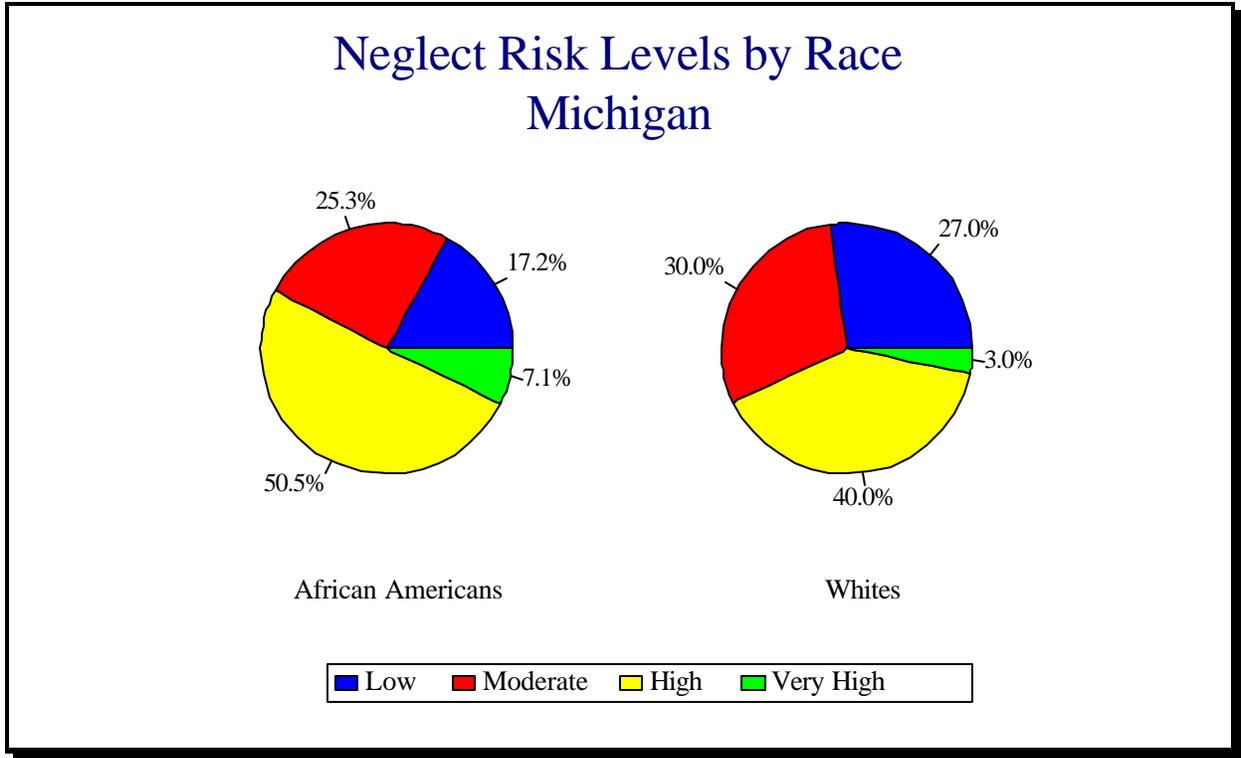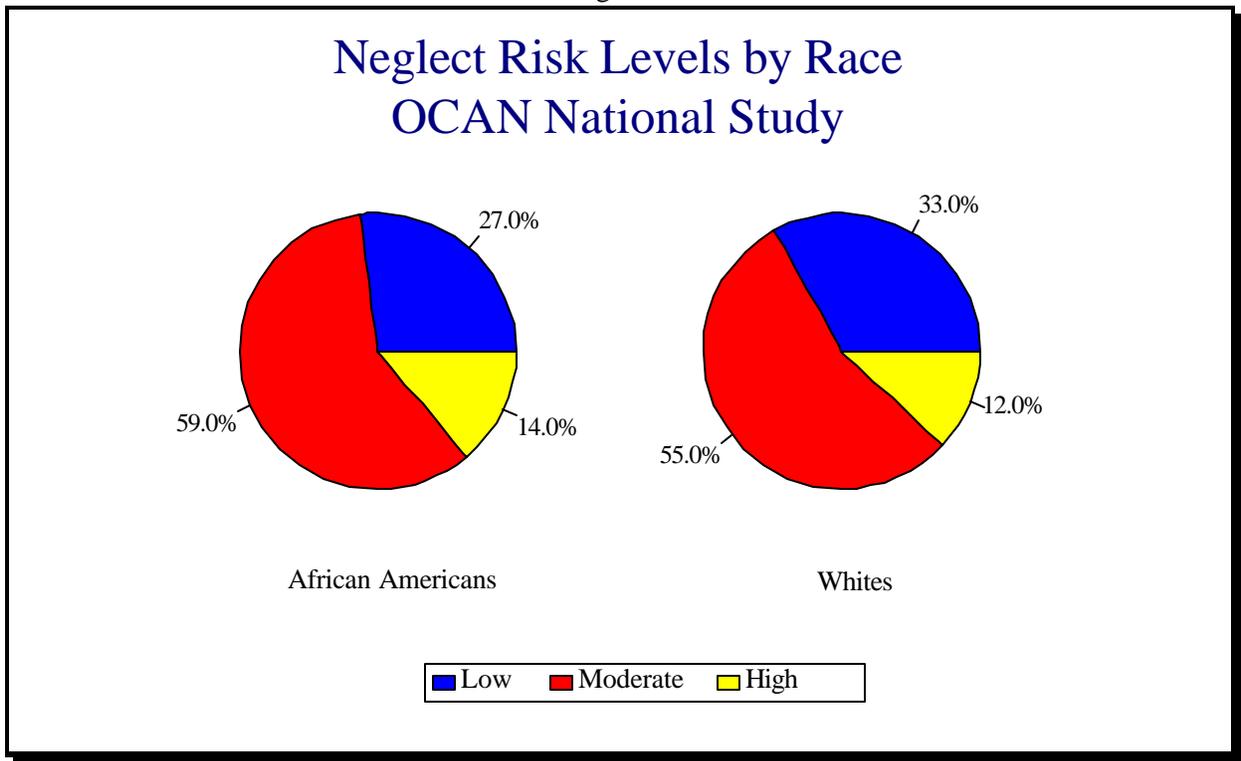12.0%

55.0%

African Americans

Whites

■ Low   ■ Moderate   □ High

Conversely, Whites as a group score higher on most items on abuse scales and have a higher average risk of abuse score than African Americans (Michigan and OCAN results are presented in Table 8). As Figures 3 and 4 illustrate, more <u>Whites</u> are rated at the higher risk levels for abuse. **When ratings from the two scales are combined by assigning the <u>highest</u> level of risk indicated by either scale, essentially equal proportions of each race are classified high, moderate, and low risk.**

	Table 8			
	Abuse Item Scores from Michigan and the National OCAN Study			
	Mean Score			
Item	**Michigan**		**OCAN**	
	Whites	**African Americans**	**Whites**	**African Americans**
Current Complaint is for Abuse	.490	.393	.553	.468
Prior Assigned Abuse Complaints	.455	.277	.451	.399
Prior CPS Service History	.465	.454	.420	.444
Number of Children in the Home	.686	.720	.716	.737
Caretaker(s) Abused as Child(ren)	.235	.141	.143	.075
Secondary Caretaker has a Current Substance Abuse Problem	.198	.152	.118	.063
Primary or Secondary Caretaker Employs Excessive and/or Inappropriate Discipline	.684	.533	.626	.526
Caretaker(s) has a History of Domestic Violence	.315	.165	.270	.165
Caretaker(s) is a Domineering Parent	.233	.157	.265	.210
Child in the Home has a Developmental Disability or History of Delinquency	.294	.167	.374	.308
Secondary Caretaker Motivated to Improve Parenting Skills	.511	.301	.336	.224

Table 8				
Abuse Item Scores from Michigan and the National OCAN Study				
Primary Caretaker Views Incident Less Seriously than Agency	.460	.478	.297	.340
Average Abuse Score	**5.026**	**3.938**	**4.569**	**3.959**

Figure 3

Abuse Risk Levels by Race
Michigan

African Americans — Low 31.0%, Moderate 40.0%, High 25.0%, Very High 4.0%

Whites — Low 18.4%, Moderate 35.0%, High 35.9%, Very High 10.7%
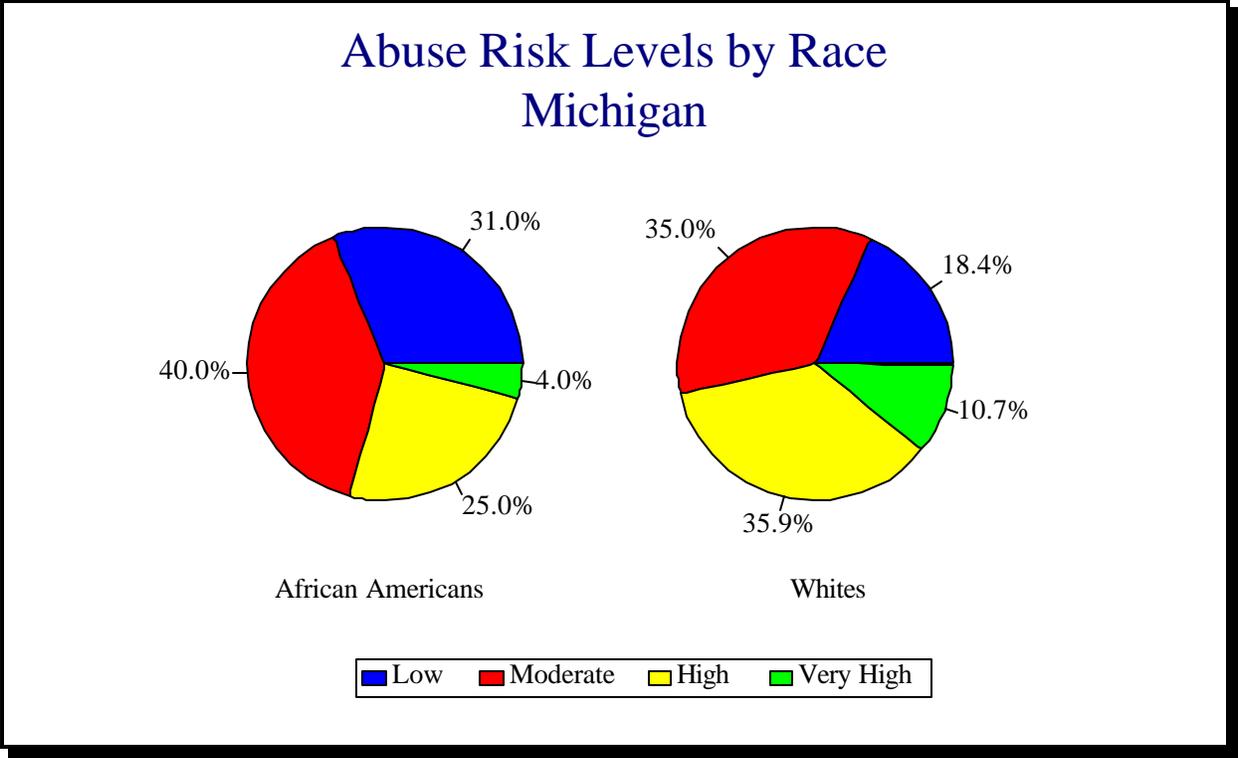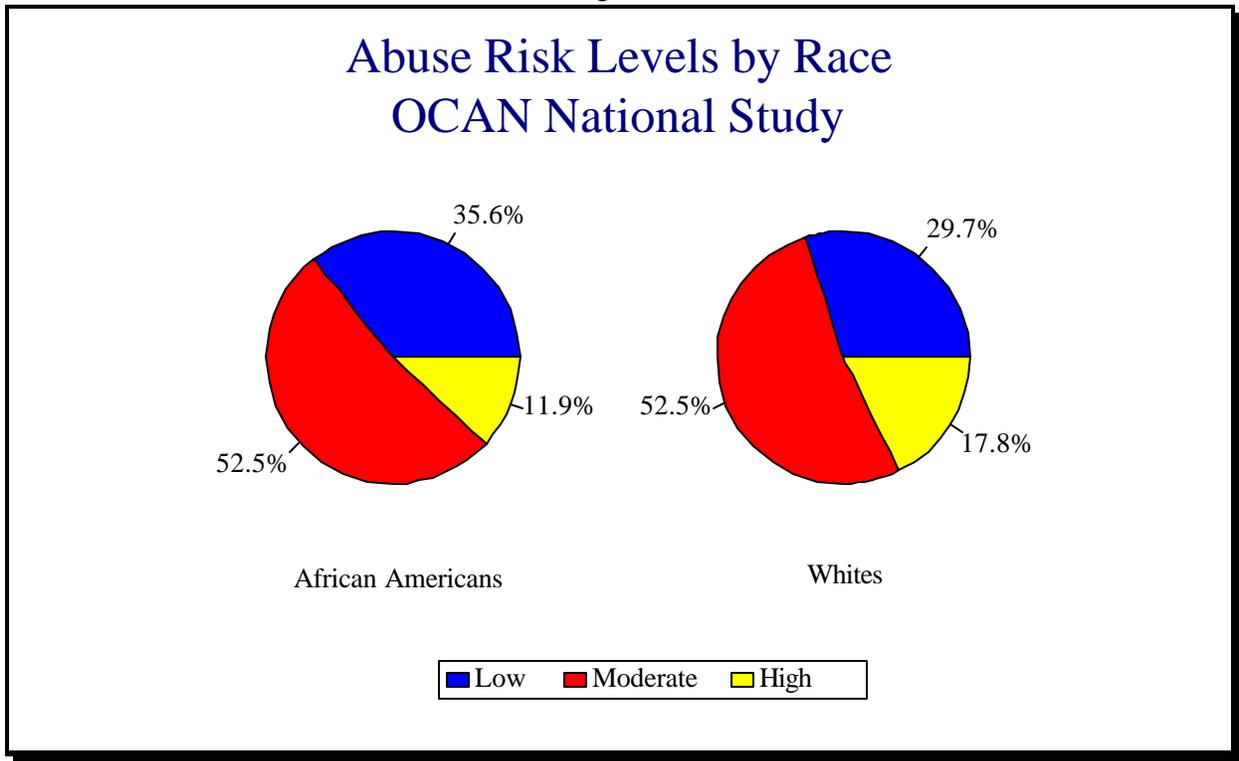
Low Moderate High Very High

Figure 4

Quite obviously, there is no racial bias produced by actuarial risk assessment systems.

The factors used to rate the risk of future abuse or neglect were selected because they relate to

subsequent reports, substantiations, injuries, and placements in families with previous investigated

allegations of abuse and/or neglect. **In instances where the average score on an item differs**

between races, this reflects actual differences in outcomes. The level of equity attained by

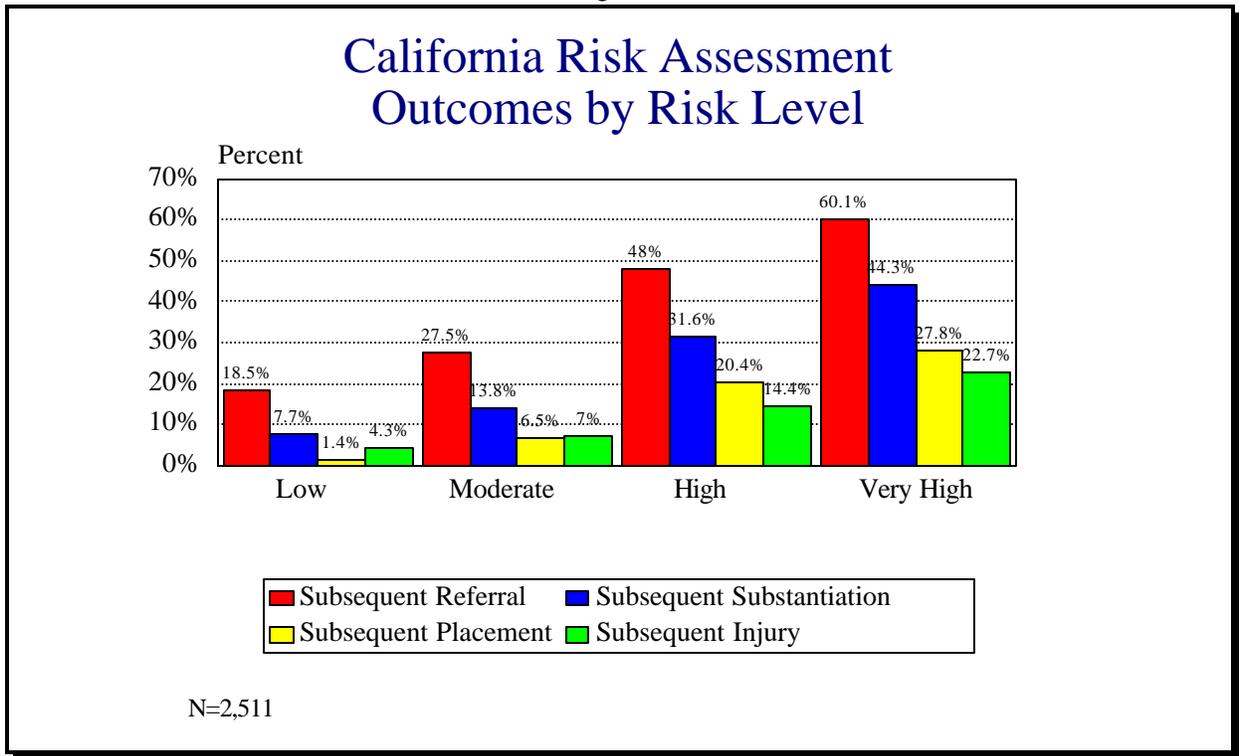actuarial risk assessment systems is rarely experienced in the human service field.

E.   IS SUBSTANTIATION A BIASED OUTCOME MEASURE?

Substantiation has been criticized as a biased outcome measure because African

Americans are more likely to have reports of abuse/neglect substantiated than Whites. While the

NIS data indicate that there is no difference in the maltreatment rates of African Americans and

Whites, the rate at which reports on African Americans are substantiated in most states is

disproportionate to their proportion of the general population. However, while the proclivity to

substantiate more reports on African Americans is a serious issue, it has little effect on risk

research. As noted earlier, risk instruments are based on <u>recurrence</u> rates. Cases included in risk

studies are only those reported to and investigated by CPS agencies. Outcomes are subsequent

reports and substantiations (and in most jurisdictions, child injuries and out-of-home placements).

The question then is: Are the disproportionate rates of substantiation for African American cases

(for <u>all</u> reports received) also found in <u>recurrence rates</u>? Based on studies conducted in nearly a

dozen jurisdictions, <u>subsequent substantiation</u> rates of the two races appear to be very similar. In

all but one CRC study there were <u>no statistically significant differences in the subsequent</u>

<u>substantiation rates of African Americans and Whites</u>. In a few studies, African American

subsequent substantiation rates were marginally higher than Whites but these differences were not

statistically significant. The OCAN study, for example, reported a subsequent substantiation rate

of 15.9% for Whites and 19.5% for African Americans (P>.05). More often, however, Whites

had marginally higher rates of subsequent substantiation. In California, the only site where

statistically significant differences in rates of subsequent substantiation were found, Whites had

the higher rate (26.1%). African Americans were re-substantiated at a 19.1% rate, while 19.7%

of Hispanics were substantiated during the two-year follow-up period. In sum, African

Americans do not appear to have a higher rate of <u>subsequent</u> substantiation than Whites.

Criticism of substantiation as an outcome measure also fails to note that actuarial risk assessment

instruments are not constructed based solely on how items relate to subsequent substantiations.

Additional outcome measures are used, including subsequent referrals, subsequent placements,

and subsequent child injuries. As Figure 5 from California demonstrates, families at each risk

level have substantially different rates reported on each outcome measure. For example, only 1.4% of all low risk cases in the California sample used to develop the scale had a subsequent out-of-home placement recorded. The placement rate increased at every risk level, to nearly 28% for cases rated very high risk.

Figure 5

California Risk Assessment
Outcomes by Risk Level

Percent

Risk Level	Subsequent Referral	Subsequent Substantiation	Subsequent Placement	Subsequent Injury
Low	18.5%	7.7%	1.4%	4.3%
Moderate	27.5%	13.8%	6.5%	7%
High	48%	31.6%	20.4%	14.4%
Very High	60.1%	44.3%	27.8%	22.7%

■ Subsequent Referral ■ Subsequent Substantiation
■ Subsequent Placement ■ Subsequent Injury

N=2,511

F.   CONCLUSIONS AND IMPLICATIONS

There is clearly an over representation of African Americans in this nation's child welfare systems, but attempts to use this fact to discourage use of actuarial risk assessment systems are misguided. Over representation of African Americans is likely due, at least in part, to unfettered discretion, not attempts to add structure to decision making. In Turning Stones: My Days and Nights with Children at Risk, Marc Parent notes, "if you get a caseworker who goes to

somebody's home and says it's fine, then it's fine. That's how important their voice is." (Parent,

1996) And, as a recent Chapin Hall report indicates, workers (and experts) frequently disagree on

what each case requires (Rossi, Schuerman, Budde, 1996). Whenever difficult decisions

affecting families are left principally to individuals, disparity is almost a certainty. The counter

claim that actuarial risk assessment systems are racially biased is based primarily on a

fundamentally flawed review of a national data base. Administrators should not be swayed by

such superficial analyses. Data presented here demonstrate that proper use of actuarial risk

assessment could well reduce over representation of minority families in the child welfare system.

To assess the potential damage that misinformation can inflict on the African American

community, child welfare administrators need only to observe what has occurred in the field of

corrections. For the last decade, groups like the Legislative Exchange Council, as well as

individuals from prominent American universities have produced reports seriously misrepresenting

facts regarding crime and incarceration. For example, these sources claimed that states that

increased the use of imprisonment the least over a ten-year period experienced the largest

increases in violent crime. This was categorically false, yet it was repeated time and again in

Congress, in nearly every state legislature in the nation, and in editorials of major newspapers

throughout the country. The results are well documented; an unprecedented increase in prison

construction and an increase in incarceration rates that has had a particularly devastating effect on

the African American community. America's policy makers are only now beginning to

acknowledge that they were misinformed.

Over representation of minorities in child welfare is a serious issue that deserves serious

study. Anything less constitutes a disservice to the field. Results from states using actuarial risk

assessment systems are compelling: these tools can help ensure that child welfare decisions are

appropriate, consistent, and equitable for all families entering the child protection system.

BIBLIOGRAPHY

National Incidence Study (NIS-3), National Center on Child Abuse and Neglect, 1996.

Parent, M., <u>Turning Stones: My Days and Nights with Children at Risk</u>. New York: Harcourt, Brace & Co., 1996.

Rossi, P., Schuerman, J., Budde, S., "Understanding Child Maltreatment Decisions and Those Who Make Them." Chapin Hall Center for Children, University of Chicago, June 1996.

Ruscio, J., "Information Integration in Child Welfare Cases: An Introduction to Statistical Decision Making," <u>Child Maltreatment</u>, Vol. 3, No. 2, pp. 143-156, May 1998.

Silver, E., Banks, S., "Calibrating the Potency of Violence Risk Classification Models: The Dispersion Index for Risk (DIFR)," Paper presented at the American Society of Criminology conference, 1998.

Silverstein, M.J., Lagios, M.D., Craig, P.H., et al, "The Van Nuys Prognostic Index for Ductal Carcinoma in Situ." <u>The Breast Journal</u> 2:38-40, 1996.

Silverstein, M.J., Lagios, M.D., "Use of Predictors of Recurrence to Plan Therapy for DCIS of the Breast," <u>Oncology</u>, Vol. II, No. 3, pp. 393-410, March 1997.

U.S. Census Bureau. 1995.

<u>Webster's II New College Dictionary</u>. Boston: Houghton Mifflin Company, 1995.