

# Sex Offender Law Report™

Vol 15 No. 5

ISSN 1529-0697

Pages 65 – 80

August/September 2014

## Sex Offender Registry and Public Access

by Roslyn Myers

Examining the effects of sex offender registry and notification laws to reduce recidivism, two studies call into question basic assumptions about the way they work. One study found that requiring sex offenders to register with police may reduce recidivism. But that study found a decrease in this positive effect when the registry information was made available to the public, suggesting that public access leads to higher overall rates of sex crime.

Analyzing a decade of data from 15 states, researchers examined separately the evolution, effects, and nature of sex offense rates as state law enforcement implemented registration and notification laws. The former require sex offenders to notify the police following their release from prison with specific information and to continue to check in at designated periods. The latter allow public access to the registry information about convicted sex offenders, nearly always as an online resource.

### Notification Laws Lower Effectiveness

The study found that registration requirements without public notification significantly reduced reported sex crime by about 13%, with larger drops in sex crime as registries grow larger. The researchers posited that the registration requirements improved police monitoring and increased the ability to locate and apprehend reoffenders, and that registry laws discouraged registered offenders from

See *PUBLIC ACCESS*, page 76

### *New Salem Witch Trials*

## Evaluating Bias in Expert Witness Conclusions of “Sexual Dangerousness,” Part II

by Daniel Kriegman, Ph.D.

**Editor’s Note:** *This is the second and final section of an article arguing that expert witnesses, who offer testimony about whether or not an offender should be civilly committed for “sexual dangerousness” may offer arbitrary opinions, which are often determinative in the court’s ruling. This article examines studies showing that the validity of expert opinions on sexual dangerousness does not meet appropriate standards of certainty for the interests at stake. (W.M. Grove, D.H. Zald, B.S. Lebow, B.E. Snitz, and C. Nelson, “Clinical Versus Mechanical Prediction: A Meta-analysis,” 12(1) Psychol. Assessment 19-30 (2000); E.S. Janus and R.A. Prentky, “Forensic Use of Actuarial Risk Assessment With Sex Offenders: Accuracy, Admissibility, and Accountability,” 40 Am. Crim. L. Rev. 1143-489 (2003); J. Monahan, Predicting Violent Behavior: An Assessment of Clinical Techniques (1981); G.G. Woodworth and J.B. Kadane, “Expert Testimony Supporting Post-sentence Civil Incarceration of Violent Sexual Offenders,” 3 L., Probability, & Risk 221-41 (2004).)*

### Clinical Bias

It is very strange—given how obvious an indicator of bias this is—that rather than adjust the base rates from the Prentky,

et al. study downward, the state’s experts appear almost always and systematically to use clinical judgment (of a kind proven to be non-predictive and to lead to errors) and then add them to misused validated factors (which, if used correctly, would improve prediction) to adjust the base rate they use upward.

The misuse of validated factors is inevitable if one uses a base rate that already includes those factors and then adjusts that rate upward using the validated factor that had already been included. For example, it is known that extra-familial child molesters have a higher rate of recidivism than the average sex offender. When such a man is being evaluated for a Section 9 hearing, it is inappropriate to utilize the Prentky statistics for an estimate of a recidivism rate for child molesters and then adjust the rate upward, because Prentky’s statistics already include the adjustment.

This prior inclusion is due to two facts. First, Prentky divided his sample into rapists and child molesters and, consistent with most other studies, reported higher rates of recidivism for molesters. This general finding leads researchers to recommend increasing the estimate of the risk for child

See *EVALUATING BIAS*, next page

### ALSO IN THIS ISSUE

Sexual Assault in Military Versus Civilian Contexts, Part II . . . . .	67
Alleged Rape and False Arrest are Serious Concern for Police and for Campus Security Procedures . . . . .	68
From the Literature . . . . .	71

**EVALUATING BIAS, from page 65**

molesters over the more general estimate of base rate for the larger group of all sex offenders. It would clearly be inappropriate to use this logic to increase the estimate of risk for a child molester over the highest base rates ever determined for child molesters just because the individual is a child molester. This would be like an insurance company putting a man into a higher risk group (to whom they charge higher rates to offset the increased risk) than a woman, because of the fact that males tend to die younger, and then—when deciding whether or not to issue the policy—deliberating, “Well, he has more male hormones than the average person and male hormones are associated with death at a younger age. Let’s move him up to the highest risk group.”

**Extra-Familial Pedophiles at Other End of Base Rate Spectrum.** Let us turn to the second fact that makes using the extra-familial factor to increase the estimate of risk almost certain to be inappropriate. A large number of child molesters commit their offenses only within their extended family. These incest offenders are known to be at the lowest risk (lower than rapists) of reoffense after a significant intervention, such as arrest and conviction and/or a period of incarceration.

Extra-familial child molesters are at the other end (higher than rapists) of the base rate spectrum. So, typically, when

evaluating men with histories of extra-familial child molesting, the state expert concludes that there is a “very high” risk of reoffense. On cross-examination, if the scientific basis for such an estimate is challenged and he is questioned on the generally low rate of recidivism for sex offenders, the expert will counter by referring to the Prentky study as showing that such estimates are probably underestimates. This shows that they are using Prentky’s higher estimates to make their risk judgment. In their reports, invariably, they then mention the extra-familial factor as leading to an “extremely high” risk (implying much more than the 50-50 which was already an overestimate of risk reported by Prentky).

**Adjustment Already Included.** The problem in using the Prentky study as an estimate of the base rate of reoffense for child molesters and then stating the man in question is at an even higher risk because he is an extra-familial child molester is due to the fact that the Prentky study already included the extra-familial factor. During the period that the men in that study were committed, the child molester group included very few (if any) men who were only incest offenders. Part of the reason they were found SD (i.e., sex offenders with a high likelihood of reoffense) to begin with—and thus were within the group that Prentky studied at the TC—is because they were extra-familial sex offenders, and thus the rate that the Prentky, et al. study produced for child

molesters already includes an adjustment for the extra-familial factor. As in the insurance example, it would be inappropriate to adjust upward a base rate determined on a group of mostly extra-familial child molesters, because the individual being evaluated is an extra-familial child molester. Yet, this occurs routinely.

I will now demonstrate that such clinical bias leads to extreme over-predictions of dangerousness.

**Data Set One: Recombitment Evaluations of Men Previously Adjudicated “Sexually Dangerous”**

In order to estimate the probability of bias, we can start by comparing the state’s Qualified Examiners’ rate of opining “sexually dangerous” in Section 9 cases (recommitment hearings of men who had been already adjudicated SD) with a reasonable estimate of the upper limit of actual recidivism for these high-risk offenders, i.e., 50%.

Again note that, for the current analysis, our use of the 50% figure makes the present argument even more compelling as it increases the likelihood that a high rate of opining “sexually dangerous” will *not* be considered biased. The reason for this is as follows: Given that Dr. 91 opined “sexually dangerous” 91% of the time (according to his own sworn testimony in court) in these cases, if the recidivism rate were 100% (all

See *EVALUATING BIAS*, page 74

**Sex Offender Law Report™**

**Editor:** Roslyn K. Myers, J.D., M.A.  
**Managing Editor:** Roslyn K. Myers, J.D., M.A.  
**Contributing Editors:** John S. Furlong, J.D., Erin Bundra, J.D.  
**Editorial Director:** Deborah J. Launer  
**Publisher:** Mark E. Peel

**Board of Advisors**

**Fred Cohen, L.L.M.**, Author, *Mentally Disordered Inmates and the Law*, Co-Editor, *Correctional Law Reporter* and *Correctional Mental Health Report*, Tucson, AZ

**Andrew Harris, Ph.D.**, Associate Dean for Research & Graduate Programs, College of Fine Arts, Humanities, and Social Sciences, University of Massachusetts Lowell

**Jill Levenson, Ph.D.**, Associate Professor, Department of Psychology & Human Services, Lynn University, Boca Raton, FL

**Stephen R. McAllister, J.D.**, University of Kansas, School of Law, Lawrence, KS

**Anita Schlank, Ph.D., L.P., ABPP (Forensic Psychology)**, Clinical Director, Virginia Center for Behavioral Rehabilitation; Editor, *The Sexual Predator*

**Phyllis Schultze, M.L.S.**, Criminal Justice Information Specialist, Criminal Justice Library, Rutgers University, Newark, NJ

**Barbara K. Schwartz, Ph.D.**, Associate, New England Forensic Associates; Former President; Author, *The Sex Offender*

**Karen J. Terry, Ph.D.**, Interim Dean of Research and Professor, John Jay College of Criminal Justice, New York, NY

**Joan Zorza, Esq.**, Founding Editor, *Domestic Violence Report*; Founding Editor, *Sexual Assault Report*, Washington, DC

Sex Offender Law Report (ISSN 1529-0697) is published bimonthly by Civic Research Institute, Inc., 4478 U.S. Route 27, P. O. Box 585, Kingston, NJ 08528. Periodicals postage pending at Kingston, NJ and at additional mailing offices. Subscriptions: \$165 per year plus postage and handling in the United States and Canada. \$30 additional per year elsewhere. Vol. 15, No. 5, August/September 2014. Copyright © 2014 by Civic Research Institute, Inc. All rights reserved. Unauthorized copying expressly prohibited. POSTMASTER: Send address changes to Civic Research Institute, Inc., P.O. Box 585, Kingston, NJ 08528. Sex Offender Law Report is a registered trademark owned by Civic Research Institute, Inc., and may not be used without express permission.

The information in this publication is not intended to replace the services of a trained legal or health, or other professional. Neither the editor, nor the contributors, nor Civic Research Institute, Inc. is engaged in rendering legal, psychological, health or other professional services. The editors, the contributors and Civic Research Institute, Inc. specifically disclaim any liability, loss or risk, personal or otherwise, which is incurred as a consequence, directly or indirectly, of the use and application of any of the contents of this publication.

For information on subscribing or other service questions contact customer service: (609) 683-4450 or order@civresearchinstitute.com

Affiliations shown for identification purposes only. Opinions expressed do not necessarily reflect the positions or policies of a writer’s agency or association.

the men were SD), he would have been right 91% of the time. If the recidivism rate were zero (none of the men were SD), he would have been right 9% of the time. Thus, because of his extremely high rate of opining “sexually dangerous,” the higher our estimate of the actual recidivism rate, the more accurate he appears. We do not have exact knowledge of the actual recidivism rate. Thus, by using an overestimate (and again, the reasons for considering this to be an overestimate were presented above) we are giving Dr. 91 the benefit of the doubt.

The other reason for using the 50% figure is that it makes the illustrative math very easy to understand, as the odds that a given offender is SD are the same as the odds of tossing a coin and having it come up heads.

**SD Determined in 47 of 49 Evaluations.** On December 7, 1992, the Justice Resource Institute listed the Section 9 (i.e., the recommitment hearing) petition status of all of the men at the TC (i.e., all of the men in the Commonwealth that were currently adjudicated “sexually dangerous”). Of the numerous men who had petitioned for a hearing, 28 had been examined by that date by one or two QEs (most by two of 11 different QEs who were doing these evaluations for the DOC at the time). These QEs comprised all of the QEs who had formed their opinion and filed a report in which they opined SD or not SD.

Of the 49 opinions filed as of that date, 47 were SD. If 50% of these men were actually likely to go out and commit a new sexual offense and were thus SD, then we should expect unbiased methods to lead to a conclusion of sexual dangerousness in about half (24 or 25) of the 49 opinions. The odds of an unbiased methodology leading to 47-or-

more-out-of-49 SD conclusions if 50% of the men were, indeed, dangerous are exactly the same as the odds of tossing a coin and getting more than 46-out-of-49 heads. The chances of that happening are one-in-500 billion. If we assume that 75% of the men at the TC petitioning for release would reoffend sexually—an assumption that is certainly false and has never been suggested by any empirical evidence—the probability of such a pattern of opining being produced by an unbiased methodology rises from one-in-500 billion to a bit more than one-in-10,000. However, when we look at individual QEs, we have larger sample sizes and can be even more certain about the presence of bias.

**Dr. 91: 91% Sexually Dangerous.** Dr. 91 is a pseudonym for a real, extremely experienced Qualified Examiner who has testified in many of these recommitment cases and was also a member of the Community Access Board (CAB) at the TC for Sexually Dangerous Persons. He has testified that he had opined “sexually dangerous” in about 82 out of 90 (or 91%) of his evaluations of men committed to the TC who were petitioning for release via a Section 9 hearing. If 50% of these men were SD, then we would expect unbiased methods to lead to a conclusion of sexual dangerousness in about half (45) of the 90 cases. The odds of an unbiased methodology leading to more than 81-out-of-90 SD conclusions if 50% of the men are, indeed, dangerous is exactly the same as the odds of tossing a coin and getting more than 81-out-of-90 heads. The chances of that happening are one-in-700-quadrillion (1:700,000,000,000,000), a number so small it has virtually no meaning.

In a murder case, a DNA blood match of one-in-seven-billion would give us virtual certainty that we had correctly identified the killer. The odds that a Qualified Examiner

could reach an 82-out-of-90 rate of opining “sexually dangerous” if his methods were not producing biased results is a number that is about a million times less likely than one-in-seven-billion. It is more than a billion times more likely that your commercial passenger jet will crash the next time you travel.

Even if we were to accept Dr. 91’s claim that the true base rate of recidivism among these men is significantly higher than 50% and we used a figure of 75%, the odds of an unbiased examiner producing an 82-out-of-90 pattern of opining SD are less than one-in-10,000. Compare that with the substantially greater odds of an average American dying in a car crash in any given year—one-in-7,000. Even if we use the extremely high figure of 75% for our estimate of recidivism in these cases, it is about one-and-one-half times (150%) more likely that you (if you are an average American) will die in a car crash this year than that Dr. 91’s methodology is unbiased.

**Dr. 98: 98% Remain Sexually Dangerous.** Dr. 98 is another real, extremely experienced qualified examiner who has testified in well over 100 of these recommitment cases and is also a member of the CAB. Dr. 98 has testified that he has evaluated 100 men committed to the TC and opined that 98 of these men remain SD. If 50% of these men are SD, then we should expect unbiased methods to lead to a conclusion of sexual dangerousness in about half (50) of the 100 cases.

The odds of an unbiased methodology leading to more than 97-out-of-100 SD conclusions if 50% of the men are, indeed, dangerous are exactly the same as the odds of tossing a coin and getting more than 97-out-of-100 heads. The chances of that

*See EVALUATING BIAS, next page*

*EVALUATING BIAS, from page 74*

happening are one-in-( $4 \times 10^{27}$ ) (1:4,000,000,000,000,000,000,000,000,000), a number so small it has no intelligible meaning; it is about a trillion billion times more likely that your commercial passenger jet will crash the next time you travel. Furthermore, if we assume that 75% of the men at the TC are SD—again, an assumption that is almost sure to be false and has never been suggested by any researcher based on actual recidivism studies—the probability of such a pattern of opining being produced by an unbiased methodology rises to one-in-five-billion. It is about 1,000 times more likely that you will die in a crash the next time you fly.

**Dr. 99: 99% Are SD.** In recent testimony, yet another member of the CAB, and an even more experienced Qualified Examiner, Dr. 99, testified that, in the course of her work on the CAB, she has carefully reviewed every one of the men that had been committed to the TC and has concluded that about 237 of about 240 men (99%) are SD. If 50% of the men there are, in fact, SD, one divided by  $10^{67}$  yields the probability that her use of an unbiased methodology could lead to such a pattern of conclusions. The denominator in that fraction is so large that there is no name for such a number. To try to comprehend how incomprehensibly large that number is, consider that  $10^{17}$  is a number greater than all the words that have ever been spoken “including all baby talk, love songs, and Congressional debates.” (E. Kasner and J. Newman, *Mathematics and the Imagination* (1940).) And  $10^{67}$  is much larger than 10 trillion  $\times 10^{17} \times 10^{17} \times 10^{17}$ . In other words, in our current analysis, the odds of such an improbable event are zero in any possible human way of imagining the meaning of such a number.

We can with absolute certainty—i.e., we can be more certain of this conclusion than of almost anything else in which we believe—be sure that an unbiased methodology did not produce this result. Furthermore, if we assume that 75% of the men at the TC are SD—once again, an assumption that is almost sure to be false—the probability of such a pattern of opining being produced by an unbiased methodology rises from zero to . . . well, a number essentially equal to zero. One divided by  $10^{27}$  is a close approximation, and  $10^{27}$  is greater than the number of all of the grains of sand on all the beaches of the world.

**Grains of Sand Required to Fill Empty Space in Universe.** We have an even more accurate estimate of the bias in the meth-

odology utilized by the experts chosen by the State to evaluate sexual dangerousness. The CAB reports have each CAB member’s conclusion regarding the sexual dangerousness of each of the 240 men at the TC. In almost every case on which I have testified in the 10 years (a good approximation of a random sample of men at the TC), the CAB has concluded unanimously that the man remains SD.

If I did not see any of the three cases on which the above-mentioned Dr. 99 voted not-sexually dangerous (not-SD) and the board was unanimous (unanimously voting not-SD, which hardly ever occurs) on those as well, the result would be 1,200 opinions by state experts at the TC each year, with about 15 being “not-SD” if the CAB unanimously agreed with Dr. 99 in those cases. If 75% of the men at the TC are SD—again, this is a number that is greater than any derived empirically and almost certain to

those sex offenders who were chosen for screening were ultimately found to be SD.

### Meaning, Determination of True Base Rate of “Sexual Dangerousness”

Since the terms “sexually dangerous person” and “likely” (to cause harm by committing sex offenses) have no independent existence in psychology or psychiatry and were created, defined, and interpreted by the legislature and the courts, we have no other source than the law and judicial interpretations to turn to for guidance about their meanings. The sexual dangerousness law has been revised and a question may be raised about whether the judicial interpretations of the older version still apply to the new law.

**Threatened Harm, Relative Certainty.** In trying to determine what the minimal degree of dangerousness is that must exist for a civil commitment of “one-day-to-life,” it has been made clear that the decisions

---

*If 50% of the men there are, in fact, SD, one divided by  $10^{67}$  yields the probability that her use of an unbiased methodology could lead to such a pattern of conclusions.*

---

be an incorrect overestimate—one divided by ( $3.5 \times 10^{125}$ ) would represent the possibility that an unbiased methodology could yield such a result. (I must also admit that I calculated the odds using the 75% figure because I could find no statistical calculator that could yield the ridiculously small probability that would result if we used the 50% figure.) The denominator of that fraction is a number greater than the number of all the subatomic particles (protons, electrons, neutrons, etc.) in the universe; it is much greater than the number of grains of sand that would be required to fill the vast volume of empty space in the entire universe.

### Data Set Two: Initial Screening, Commitment Evaluations of Men Proposed as Possibly “Sexually Dangerous”

Based on the data collected by Dr. Robert Prentky and his research staff, over the 25-year period from 1959 to 1984, 5,000 sex offenders were screened; probable cause was found for 2,000 who were subsequently sent for 60 days of observation. (Prentky, personal communication, 2002.) Five hundred of those 2,000 (25%) were ultimately found to be SD by the superior court judges of the Commonwealth of Massachusetts, or 500 of the original group of 5,000 screened men (10%). Thus, 10% of

under the old law must be used for guidance. This is so because, first of all, the old version was less restrictive (there were fewer criteria that had to be met in order to conclude that a man was “sexually dangerous”) as it did not require a finding of a mental disorder that causes a serious difficulty in controlling behavior. It merely required repetitive sexual misconduct of the type that, from a legal standpoint, is present in most of the cases considered under either version of the law. Thus, comparing an offender to a known standard established by the superior courts in the past should provide a valid yardstick to assess the minimal mixture of the seriousness of the predicted type of offense (“the threatened harm”) and the “relative certainty of the anticipated harm” that are usually the essential factors in assessing the “risk of reoffending” and determining what is “likely” as defined by the law.

*In assessing the risk of reoffending, it is for the fact finder to determine what is “likely.”* Such a determination must be made on a case-by-case basis, by analyzing a number of factors, including *the seriousness* of the threatened harm, *the relative certainty* of the anticipated harm, and the possibility of

See *EVALUATING BIAS, next page*

---

*EVALUATING BIAS, from page 75*

---

successful intervention to prevent that harm. (**Commonwealth v. Boucher**, supra, emphasis added.)

**Definition of “Likely.”** Secondly, beyond the reasonableness of using the old standard of a less restrictive definition of “sexual dangerousness” to determine the minimal conditions that must be met, we have been specifically instructed by the Supreme Judicial Court of Massachusetts to use this very yardstick, i.e., to rely on past judicial interpretations of the terms under the old law. This standard—what the courts decided about the meaning of “likely”—remains valid, since the meaning of “likely” has not changed from the old to the new statute:

Further evidence of legislative intent can be gleaned from looking at how the term “likely” was interpreted in G.L. c. 123A before that statute was amended in 1999. St.1999, c. 74, § 6. . .Because the use of “likely” in the current and prior versions of G.L. c. 123A, § 1, is so similar, and given the well-settled and logical rule of statutory construction that, when the Legislature re-enacts a statute using the same language, the Legislature must be presumed to have adopted the prior judicial construction of that language, we can look for guidance to

judicial interpretations of the term in its prior incarnation. (Id.)

Rather than use each expert’s own subjective standard of what “likely” means (what “seriousness of the threatened harm” combined with what degree “of relative certainty” is sufficient cause for deprivation of liberty), I am simply pointing out that any expert opinion on a matter that is legally defined must follow the legal construction of the terms. Since the application of the terms to a specific case is determined by the fact finder, then—rather than having the expert instruct the fact finder on the meaning of “likely”—the expert must be bound to use the term *as the fact finders* have done in the past. And the fact finders who set the existing standard were the superior court judges in their judicial interpretations of the law (as guided by the higher courts). An expert can only present a valid opinion (as an expert in the field of dealing with men who had been adjudicated SD) about whether a man is SD by using the yardstick that was derived from what the judicial fact finders have concluded in the past, not derived from what the particular expert believes is an acceptable or unacceptable level of risk.

**Smaller Percentage Appropriately Found SD Today.** Starting with this standard established by the superior court judges—10% of screened offenders and 25% of offenders temporarily committed after screening and a finding of probable

cause were deemed SD—we know that a smaller percentage of men could be appropriately found SD today. This is so because of the following:

1. Less dangerous men are now being selected and included in an expanded group of sex offenders who are chosen for screening. Under the old law, 10% of screened offenders were found to be SD and only judges and prison superintendents referred sex offenders for screening. Under the new version of the law, elected officials (DAs) select who gets screened. The result is that, operating with a high degree of sensitivity to possible political repercussions if any sex offender who is released commits another offense, some DAs refer nearly every sex offender and others refer many who are grossly inappropriate and who would never have been screened in the past.
2. In the past, the initial screeners found less of the men they screened to be possibly SD (well under 50%) in comparison to Dr. Meadowsquire’s (a pseudonym) findings (presented below), in which she found 65% of the men to be SD at initial screening. With the current QEs’ overall SD rates at initial screening averaging over well over 50%, a much higher percentage of less dangerous men are now temporarily committed for 60

*See EVALUATING BIAS, next page*

*EVALUATING BIAS, from page 76*

days for evaluation and a subsequent commitment trial.

3. The criteria are now more stringent in that they now require the additional proof of the existence of a mental disorder that leads to a serious difficulty in controlling one's sexual behavior.

Thus the 10% figure from the past occurred when more dangerous men were initially screened and a smaller percentage of them were passed along through the screening process and the bar was lower for a finding of SD. So, today, less than 10% of those selected for screening should be found SD and less than 25% of those committed for a 60-day evaluation should be found SD.

However, again giving the QEs the benefit of the doubt, we can use these certain overestimates of risk as a comparison base rate to evaluate the presence or absence of bias.

### **Does QE Show Evidence of Bias When Performing Screening Evaluations at Probable Cause Stage?**

Based upon the way the superior court judges determined sexual dangerousness, and based on the considerations and the numbers presented above, the maximum percentage of screened men who today could reasonably be found SD in accordance with the proper interpretation (judicial construction) of the law is 10%. We can compare Dr. Meadowsquire's rate (according to her testimony June 23 2004, in which she said she found 65% of the 40 men she had screened to be SD) to the 10% rate the judges ultimately found when judges were the fact finders. (Note that Dr. Meadowsquire was typical of the Qualified Examiners and is just being used as an example for purposes of this analysis.)

**Odds if Judges Were Reasonable.** If the judges were more or less reasonable, then the true base rate is at most—because of the three considerations above—10% of the men seen at the screening/probable cause level reasonably meeting the legal criteria for being found SD. Meadowsquire found 65% (of 40) to be SD—then the odds of an unbiased examiner happening to be asked to evaluate a group of 40 sex offenders in which 26 or more could reasonably be found to be SD can be calculated. These odds are  $< 0.000000000000000006$  or  $< 6/100,000,000,000,000,000$ , or much less than one-in-10-quadrillion.

**Odds if Judges Were Wildly Irresponsible.** For argument's sake, we could also

assume the judges were wildly irresponsible and missed as many as half of the men who met the legal criteria for being considered SD. If this assumption about the judges were true, we could then compare Meadowsquire's rate to a more accurate base rate of 20%. (We simply double the 10% rate to make up for the fact that the judges were so wildly irresponsible and off by such a wide margin.)

If the judges were wildly irresponsible, then a more accurate base rate might be double what the judges found, or a rate as high as 20% of the men seen at the screening/probable cause level could reasonably meet the legal criteria for being found SD. Meadowsquire found 65% (of 40) to be SD—then the odds of an unbiased examiner randomly being asked to evaluate a group of 40 sex offenders in which 26 or more could reasonably be found to be SD can be calculated. These odds are  $< 0.000000000785$  or  $< 1/1,000,000,000$  or much less than one-in-a-billion.

### **Does QE Show Evidence of Bias When Performing Post-Screening (60-Day Commitment) Evaluations at Trial Level?**

According to the superior court judges' determinations of sexual dangerousness, based on the considerations and the numbers presented above, the maximum number of men committed for 60-days observation (probable cause had been found) who could meet the legal criteria for being considered SD is 500 out of 2,000+ or  $< 25\%$ . Since, as noted above, less dangerous men are being referred by elected officials for possible commitment, and since a higher percentage are being considered dangerous at the screening/probable cause level, and since more stringent criteria have to be met for a determination of SD under the new law, we can be certain that 25% is an overestimate of the number of men who would meet the legal criteria today at the initial commitment trial stage. Thus, giving Dr. Meadowsquire the benefit of the doubt, we can compare her rate (according to her testimony) of finding 80% of the 55 men she has evaluated at this level to be SD to the judges' rate of 25%.

**Odds if Judges Were Reasonable.** If the judges were more or less reasonable, then the base rate today would be, at most, 25% of the men seen at the commitment trial level could reasonably meet the legal criteria for being found SD. Meadowsquire found 80% (of 55) to be SD—then the odds of an unbiased examiner happening to be asked to evaluate a group of 55 sex offenders in which 44 or more could reasonably be found to be SD can be calculated. These odds are  $< 0.00000000000000002$  or  $<$

$1/50,000,000,000,000,000$  or less than one-in-50-quadrillion.

**Odds if Judges Were Wildly Irresponsible.** We could also assume the judges were wildly irresponsible and, instead of finding only one-quarter of the men SD, should have found that more than one-third met the legal criteria for being considered SD. We could then compare Meadowsquire's rate to a more accurate base rate of 35%.

If the judges were wildly irresponsible and a more accurate base rate is 35% of the men seen at the commitment trial level could reasonably meet the legal criteria for being found SD, and Meadowsquire found 80% (of 55) to be SD, then the odds of an unbiased examiner happening to be asked to evaluate a group of 55 sex offenders in which 44 or more could reasonably be found to be SD can be calculated. These odds are  $< 0.0000000001$  or  $< 1/100,000,000,000$  or less than one-in-100-billion.

**Odds if Judges Were Totally Out-to-Lunch.** We could also assume the judges were "totally out-to-lunch" in their interpretation of the statute and, instead of finding only one-quarter of the men SD, should have found that twice as many men (or 50%) met the legal criteria for being considered SD at the trial level. We could then compare Meadowsquire's rate to a more accurate base rate of 50%.

If the judges were "totally out-to-lunch" and a more accurate rate is double what the judges found, then 50% of the men seen at the commitment trial level could be expected to meet the legal criteria for being found SD, and Meadowsquire found 80% (of 55) to be SD, then the odds of an unbiased examiner happening to be asked to evaluate a group of 55 sex offenders in which 44 or more could reasonably be found to be SD can be calculated. These odds are  $< 0.000005$  or  $< 5/1,000,000$  or less than one-in-200,000.

### **Can We Conclude That QE Is Biased?**

Let us first only consider the most favorable (for her) estimate of Meadowsquire's bias, the one produced in looking at her SD conclusions at the trial stage and comparing them with twice the actual rate of SD (i.e., twice the rate determined to exist by the superior court judges and thus twice the de facto definition of the legal concept of sexual dangerousness as established by the courts). In that analysis, there is less than a one-in-200,000 chance that she just happened to be asked to evaluate a group of

*See EVALUATING BIAS, next page*

*EVALUATING BIAS, from page 77*

sex offenders who reasonably could be concluded to have evidenced as high a degree of dangerousness as she opined. Remember, this most favorable estimate is based on the following unlikely assumption: The judges were “totally out to lunch” in their interpretation of the law and should have committed twice the number they committed.

Given that assumption, there is less than a one-in-200,000 chance that Dr. Meadowsquire is an unbiased, reasonable examiner who just happened to get a group of 55 offenders who had an unusually high rate of dangerousness. So, how likely is one-in-200,000? Well, as noted, the odds are one-in-7,000 that you, the reader, if you are an average American, will die in a car crash this year. So, you are almost 30 times more likely to die in a car crash this year compared to the likelihood that Meadowsquire is unbiased (if the judges were “totally out to lunch,” that is, and should have decided SD at twice the rate they did).

Over a lifetime, the odds are one-in-100 that you will die in a car crash. Since dying in a car crash is a relatively unusual way to die, we can begin to see that one-in-100 means an event is fairly rare, not to mention one-in-200,000. Yet, you are 2,000 times more likely to die in a car crash compared to the likelihood that Meadowsquire is unbiased (if the judges were “totally out to lunch” and should have decided SD at twice the rate they did). This is a stronger finding than virtually anything in the social sciences and is far, far stronger than anything we have learned about predicting sex offense recidivism.

Finally, if the judges were more or less reasonable, you are over 10 billion times more likely to die in a plane crash the very next time you fly on a regularly scheduled commercial passenger jet compared to the likelihood that Meadowsquire is unbiased.

### **Another Demonstration of State Examiners' Extreme Bias**

Since sexual dangerousness is a legal concept that is defined by the legislature and interpreted by the courts, we have another way of knowing that Meadowsquire's SD opinions are biased beyond any possibility of doubt. Let us again use twice the true rate of sexual dangerousness at the initial commitment hearings—50% instead of the 25% that was established by judicial practice over 30 years—and take a look at what that would mean if Meadowsquire were right and were correctly determining sexual dangerousness.

Though far less than Meadowsquire's rate of finding men SD, just this doubled rate of 50% would ultimately bankrupt the Commonwealth. This can be considered another indication of extreme bias, as we must assume that the legislature did not intend for this statute to incarcerate so many sex offenders beyond their criminal sentences that the state would be completely bankrupted.

The cost of actually committing those whom Dr. Meadowsquire considers SD would eventually make the primary economic activity of the Commonwealth of Massachusetts the commitment and treatment of the sexually dangerous. It would require the construction of a new 250-bed prison every three years for the next 40 years (at which point the release rate and death from age and disease would begin to match the commitment rate). The annual cost, in today's dollars, to confine and treat the SD would then be more than \$500 million. If the men, who are entitled to petition for a review hearing every year, actually received a hearing once in every three years (as they do now), it would also require 1,000 additional recommitment trials annually (almost always tried before a jury and lasting a minimum of four days and, in some cases, as much as four weeks), necessitating more than 27 new, full-time superior court judges and their staff. Overall, such a commitment rate would lead to almost a 50% increase of the number of prison inmates in Massachusetts.

### **Beyond Meadowsquire: Is There Pattern of Bias Among Qualified Examiners?**

Based on his testimony, Dr. Goldsmith (another pseudonym) opined that 35 out of 57 men that he screened (61%) met the criteria for being considered SD. Remember that only 10% (500) of the 5,000 men referred for screening by a QE were ultimately found to be SD. The odds that an unbiased method could produce a finding of 35 or more out of 57 men are SD at the screening level, given the established 10% rate for this legally determined classification, are  $< 3.7 \times 10^{21}$  (less than 1/400,000,000,000,000,000,000) or less than one-in-400-quintillion. If the judges were wildly irresponsible and a more accurate base rate were 20%, the odds of Dr. Goldsmith just happening to be assigned to a group of 57 men, of whom 35 or more could reasonably meet the criteria for being SD, would be about 1/10,000,000,000 or one-in-10-billion.

Other experts for the state have similar base rates ranging (again, based on their own estimates during testimony) from opining SD in 40% to as many as 70% of the cases seen at the probable cause stage. The odds against

their being unbiased are just as high, even when one considers the much lower rate of opining SD that some have (possibly as low as 40%), once the total number of cases gets fairly high. For example, for an examiner who found 40 out of 100 men to be SD (a 40% rate), even if we assumed the judges were wildly irresponsible and that an accurate base rate is 20%, the odds of such an examiner being unbiased would be less than one-in-400,000. After a couple of years of extra cases, when the examiner has seen 150 men, at the same rate of opining SD, the odds of being unbiased would be less than one-in-50-million. The point is that any examiner who has evaluated more than a handful of men and whose base rate significantly exceeds the 10% rate established by the superior court judges can be known to be biased.

### **Without Bias, Significant Deviations From Base Rate Become Rare When We Have Good Deal of Experience (When *n* Is Large)**

The base rate for an unbiased coin is 50% heads and 50% tails. Yet, we can never see the base rate on any one toss of three unbiased coins; with three coins, they can never come up half heads and half tails. The odds are exactly 50% that, if one tosses three unbiased coins, we will see two or more heads (and 50% that we will see two or more tails). Heads coming up two or more times on a three-coin toss is thus an ordinary event that is as likely to happen as not.

**Tossing for 15 Billion Years.** However, if we toss six unbiased coins at once, the same ratio of 2/3 (i.e., four) or more heads becomes less likely; it will happen only 33% of the time. If we toss 30 unbiased coins at once, it becomes unlikely that we will see 20 (i.e., 2/3) or more heads; with 30 coins, getting 2/3 or more heads will occur less than 5% of the time. At 300 coins, 2/3 (i.e., 200) or more heads, for all practical purposes, becomes a statistical impossibility; on average, it will happen once in 250 million tossings of 300 coins.

If an army of experimenters—who, every second, could toss 1,000 coins, tabulate the result, and gather the coins for another toss—had begun at the Big Bang (15 billion years ago) and continued until today, they would never have seen 667 (i.e., 2/3) or more heads in any unbiased 1,000-coin toss. If the universe started over and ran again for another 15 billion years and they kept tossing and counting, they still would not see 667 heads come up. Indeed, the odds are less than one-

*See EVALUATING BIAS, next page*

**EVALUATING BIAS, from page 78**

in-100-million—if we could toss the coins, tabulate the results, and gather them up for another toss once every second for 15 billion years—that we would ever see 667 or more heads out of 1,000 coin tosses.

**Dictating New Standard of Own Choosing Should Not Be Allowed.** Thus, when we have numerous iterations, significant, unbiased deviation from an established base rate become highly unlikely. If experts are interpreting the data correctly and using a proper understanding of the legally defined terms—in which it is for the fact finder to determine what level of risk is “likely” to lead to serious harm by considering (1) the seriousness of the likely harm, (2) the certainty of the harm, and (3) interventions that could reduce the risk—then the experts must follow the fact finders who, under the supervision of the statute and the courts, over a period of 30 years, established the meaning of sexual dangerousness. The experts should not be allowed to dictate a new standard of their own choosing to the fact finders of today. Over the three decades from 1959 until 1990, the fact finders (the superior court judges) established a clear standard in which 10% (i.e., 600 of the approximately 6,000 men who were referred for screening) of the sex offenders considered were ultimately deemed to meet the criteria.

If the entire group of state experts were pooled together and we looked at their average

rate of opining SD at the probable cause level during the period of this study (which would be well over 40%, the lowest rate reported) we would then be looking at over 500 cases. With such a large *n*, the probability of obtaining such a high average rate of SD opinions without the presence of extreme bias becomes absurdly infinitesimal.

**No Doubt of Biased Methodological Approach.** If invoking public safety, one chooses to err on the side of over-committing sex offenders who may not meet the criteria for being considered SD, one can assume the judges were too cautious in assessing dangerousness that would justify an indefinite loss of liberty and that they should have committed twice as many men. Then we should only consider the possibility of bias if the examiner opines SD at more than twice the judges’ base rate (i.e., if the examiner’s rate significantly exceeds 20%).

Even with these assumptions, if the examiner’s SD rate at probable cause is above 35% and more than 22 evaluations have been done, the indication of bias is as “statistically significant” at the standard for findings in the scientific literature. By raising either number (the examiner’s SD rate or the number of evaluations performed), the result starts to be “highly significant.” When you get into the realm of actual Qualified Examiners’ patterns of opining (40% to 70% SD findings at the initial screening stage and typically more than 50 to 100 cases), the results become astronomically unintelligible, requiring the

use of numbers that have no real meaning to us (quadrillion, quintillion, etc.).

The odds are so small that an unbiased examiner could reasonably conclude that such a high percentage of offenders meet the legal criteria for being found SD that we can state beyond the possibility of doubt that a methodological approach that yields highly biased, inaccurate results must be operating.

*Dr. Daniel Kriegman, a licensed psychologist, served as the Director of Supervision and Training and the Director of Intake and Treatment Planning at the Massachusetts Treatment Center for Sexually Dangerous Persons at the Massachusetts Correctional Institution at Bridgewater, where he trained as a student and later functioned as a clinical leader. He is a Qualified Examiner under the Sexual Dangerousness statute (M.G.L. Chapter 123A). He formed the Human Services Cooperative. He is the co-author (along with Dr. Malcolm Slavin) of the acclaimed book, The Adaptive Design of the Human Psyche: Psychoanalysis, Evolutionary Biology, and the Therapeutic Process (1992, Guilford Press), that created the psychoanalytic paradigm known as “evolutionary psychoanalysis,” and co-editor (with J.G. Teicholz) of Trauma, Repetition, & Affect Regulation: The Work of Paul Russell (1998, The Other Press). In addition, he has published widely on topics related to the evolutionary understanding of human behavior and the theory and practice of depth psychological (psychoanalytic) approaches to psychotherapy. Dr. Kriegman has been on the Faculty of the Massachusetts Institute for Psychoanalysis, and was a founding board member of the Psychoanalytic Couple and Family Institute of New England, Inc.*

**Table 1: Summary Chart Evaluating Bias in Expert Witness Conclusions of “Sexual Dangerousness”**

Examiner	Rate of Opining Sexually Dangerous	If True Base Rate Is	Odds That an Examiner is Using an Unbiased Methodology
Dr. 91 is one of the most experienced QEs in the state. He was one of the chairpersons of the Community Access Board (CAB) that evaluates each of the men at the Treatment Center on an annual basis and determines whether they remain sexually dangerous.	82 out of 90 MGL c. 123A § 9 cases (91% SD)	0.4 0.5 0.75	Less than 1 ÷ 10 <sup>23</sup> Less than 1 ÷ 10 <sup>17</sup> Less than 1 in 10,000
Dr. 99 is as experienced as Dr. 91 and for years was a member and chairperson of the CAB.	237 out of 240 cases reviewed (99% SD)	0.4 0.5 0.75	Less than 1 ÷ 10 <sup>87</sup> Less than 1 ÷ 10 <sup>65</sup> Less than 1 ÷ 10 <sup>25</sup>
The members of the CAB that evaluates an offender’s progress in treatment and decides whether he remains sexually dangerous.	~1,185 out of 1,200 opinions (99% SD)	0.4 0.5 0.75	Incalculably small Incalculably small Less than 1 ÷ 10 <sup>124</sup>

**Notes:**

- 1 in 10,000 is a number 150% less likely than you (if you are an average American) dying in a car crash this year.
- 10<sup>17</sup> is a number greater than all the words ever spoken, including “baby talk, love songs, and all Congressional debates.”
- 10<sup>25</sup> is a number greater than all the grains of sand on all the beaches of the world.
- 10<sup>65</sup> is *much* greater than 10<sup>17</sup> cubed, or Ten Trillion times AWES times AWES times AWES.
- 10<sup>124</sup> is a number greater than the number of all the subatomic particles (protons, electrons, neutrons, etc.) in the universe; it is *much* greater than the number of grains of sand that would be required to fill the vast volume of empty space in the entire universe. ■





## **Authorized Electronic Copy**

This electronic copy was prepared for and is authorized solely for the use of the purchaser/subscriber. This material may not be photocopied, e-mailed, or otherwise reproduced or distributed without permission, and any such reproduction or redistribution is a violation of copyright law.

For permissions, contact the **Copyright Clearance Center** at  
<http://www.copyright.com/>

You may also fax your request to 1-978-646-8700 or contact CCC with your permission request via email at [info@copyright.com](mailto:info@copyright.com). If you have any questions or concerns about this process you can reach a customer relations representative at 1-978-646-2600 from the hours of 8:00 - 5:30 eastern time.