



Agreement between court decisions and evaluator opinions in SVP trials

Rachel Kahn, Ph.D. & Sharon Kelley, Psy.D.

ATSA: Kansas City, MO

October 26, 2017

What do courts think of expert witnesses?

- Edens et al. (2012): Examined legal cases published in the past 20 years for evidence of disparaging remarks towards psychologists and psychiatrists (i.e., bias; charlatan).
 - N = 160 cases with 245 derogatory statements from attorneys and judges.
- Statements included:
 - being for sale (28%)
 - partisan/advocate (27%)
 - biased (21%)
 - pseudoscience (14%)
 - mysticism (6%)

What do jurors think of expert witnesses?

- Perceived impartiality, credibility, clarity and efficacy are frequently cited as being most influential (Brodsky et al., 2010; Cramer et al., 2010; Freckelton et al., 1999; Young et al., 1999)



What do jurors think of expert witnesses?

- Cramer, Parrott, Gardner, Stroud, Boccaccini, & Griffin (2014)
 - N = 324 mock jurors rated experts seen in a trial via video recording
 - Credibility – Witness Credibility Scale
 - Measures likeability, trustworthiness, confidence, & knowledge
 - Efficacy – Observed Witness Efficacy Scale
 - Measures poise and communication style
 - Personality – Five-Factor Mini-Markers Scale
 - Measures neuroticism, extraversion, openness, agreeableness, & conscientiousness



What do jurors think of expert witnesses?

- Cramer, Parrott, Gardner, Stroud, Boccaccini, & Griffin (2014)
 - Character but not efficacy was associated with jurors agreement with the expert on sentencing
 - Character was correlated with
 - Openness
 - Low neuroticism
 - Agreeableness
 - Conscientiousness
 - Likeability
 - Trustworthiness
 - Knowledge
 - Confidence

What do jurors think of SVP evaluators?

- Boccaccini et al. (2014)
 - Surveyed actual jurors (N = 161) from 14 SVP trials in Texas
- When trials had both a state and defense expert, most jurors believed
 - Testimony from the experts was important (98%)
 - Experts were either OK (42%) or Very Good (56%) at identifying people who are high risk
 - Actuarial measures help experts make accurate decisions (67%)*
 - The experts were honest (74%)
 - The reason the experts did not agree was because one expert did a better job than another (50%) and because it was a difficult and complicated case (34%)

How do jurors make decisions in SVP trials?

- Mock juror studies –
 - More influenced by unstructured clinical expert testimony vs. actuarial expert testimony (Krauss et al., 2011; McCabe et al., 2010)
 - Mock jurors credit actuarial results when it confirms perception of risk and discredits actuarial results when it indicates low risk (Scurich & Krauss, 2013)
 - Most influenced by past offenses (Krauss et al., 2011)
 - Being referred for SVP may be sufficient grounds for commitment for most mock jurors (Scurich & Krauss, 2014)
 - Female jurors more strongly in favor of civil commitment (McCabe et al., 2010), especially when respondent is described as “a psychopath” (Guy & Edens, 2006)




How do jurors make decisions in SVP trials?

- ▶ Turner et al. (2015): surveyed actual jurors (N = 462) across 40 SVP trials in Texas
- ▶ Most influenced by:
 - ▶ Offense history
 - ▶ Diagnosis
 - ▶ Failure to take opportunities to change and lack of remorse
 - ▶ Clinically based testimony (vs. actuarial results)
 - ▶ Importance of PCL-R increased with higher scores
 - ▶ Female jurors more influenced by PCL-R rating
- ▶ Note: in Texas SVP trials, respondent has no 5th Amendment right in civil trials and must testify in front of a jury; trials are only done in one county; all initial SVP trials resulted in commitment since law enacted in 1999



The current study

- Exploratory in nature
 - How often do the courts agree with a “neutral” evaluator in Wisconsin SVP trials?
 - Do jury trials and bench trials have a differing rate of agreement?
 - Do evaluator recommendations predict court outcome after controlling for other variables?
- 



Some context

- In Wisconsin, individuals undergo two separate risk assessments to determine their need for commitment as SVP: a pre-probable cause evaluation and a post-probable cause evaluation.
- Both evaluators testify at the pre-commitment hearings
- Following commitment, an annual risk assessment to determine their continued need for commitment is required (post-commitment evaluation)
 - When respondent petitions for release, the evaluator who completed the post-commitment report testifies



More context

- This study focuses on the following evaluations:
 - The initial post-probable cause evaluation (.04) used in the pre- commitment trial
 - The annual report to the court (.07) used in the post-commitment trial
 - The evaluator is assigned to these cases at random as part of their state employment and is considered a “neutral” evaluator.
 - Either side can retain additional expert witnesses (they are not considered “neutral” for the purposes of this study).
- The threshold for determining an individual’s need for commitment is “more likely than not” to commit a sexually violent act, which has been interpreted by case law to mean over 50%.



Current Study

Trial Descriptives

- N = 214 court outcomes from 2012 – 2016
 - $n = 26$ were excluded
 - 10 cases discharged due to death of patient
 - 13 cases dismissed or stipulated prior to evaluator opinion
 - 2 cases stipulated where evaluator opinion was equivocal
 - 1 mistrial

FINAL SAMPLE: $n = 188$ trials

Patient Descriptives

- Patients were all adult male sex offenders residing at SRSTC or on Supervised Release (SR).
- Ages 24 to 80 ($M = 51.54$, $SD = 11.16$)
- Ethnicity: 61.1% White/Caucasian; 31.1% Black; 6.1% Native American; 1.6% Other
- Static99R Score: 1 – 9 ($M = 5.38$, $SD = 1.71$)
- PCL-R Score: 9 – 37.5 ($M = 24.46$, $SD = 5.31$)

DHS Evaluator Descriptives

- $N = 15$
- 67% male / 33% female
- Number of trial cases
 - Range: 1 – 28 cases
 - $M (SD) = 13.0 (8.24)$

Description of Main Analyses

- Agreement: *Kappa* calculation of agreement rate between evaluators and courts and adjusts for chance rate of agreement
 - Ranges from -1 to 1 (< 0 equivalent to less than chance agreement, 0 being no agreement, and 1 equivalent to perfect agreement)
- Disagreement: McNemar tests comparing paired proportions in evaluator recommendation and court decision.
- Logistic regression: outcome is prediction of court decision

Interpretation of Kappa

Table 2

Interpretation of Kappa

Poor Slight Fair Moderate Substantial Almost perfect

Kappa 0.0 .20 .40 .60 .80 1.0

Kappa Agreement

< 0 Less than chance agreement

0.01–0.20 Slight agreement

0.21– 0.40 Fair agreement

0.41–0.60 Moderate agreement

0.61–0.80 Substantial agreement

0.81–0.99 Almost perfect agreement

Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159-74.



Case Descriptives

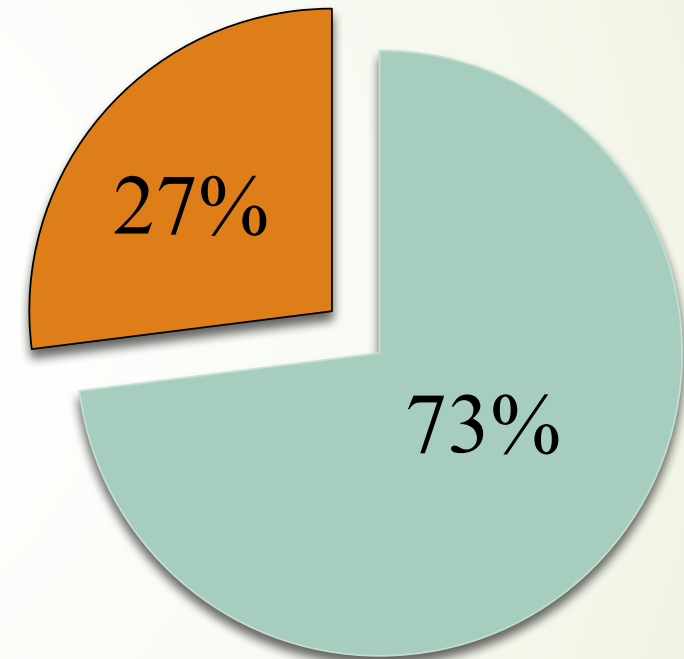
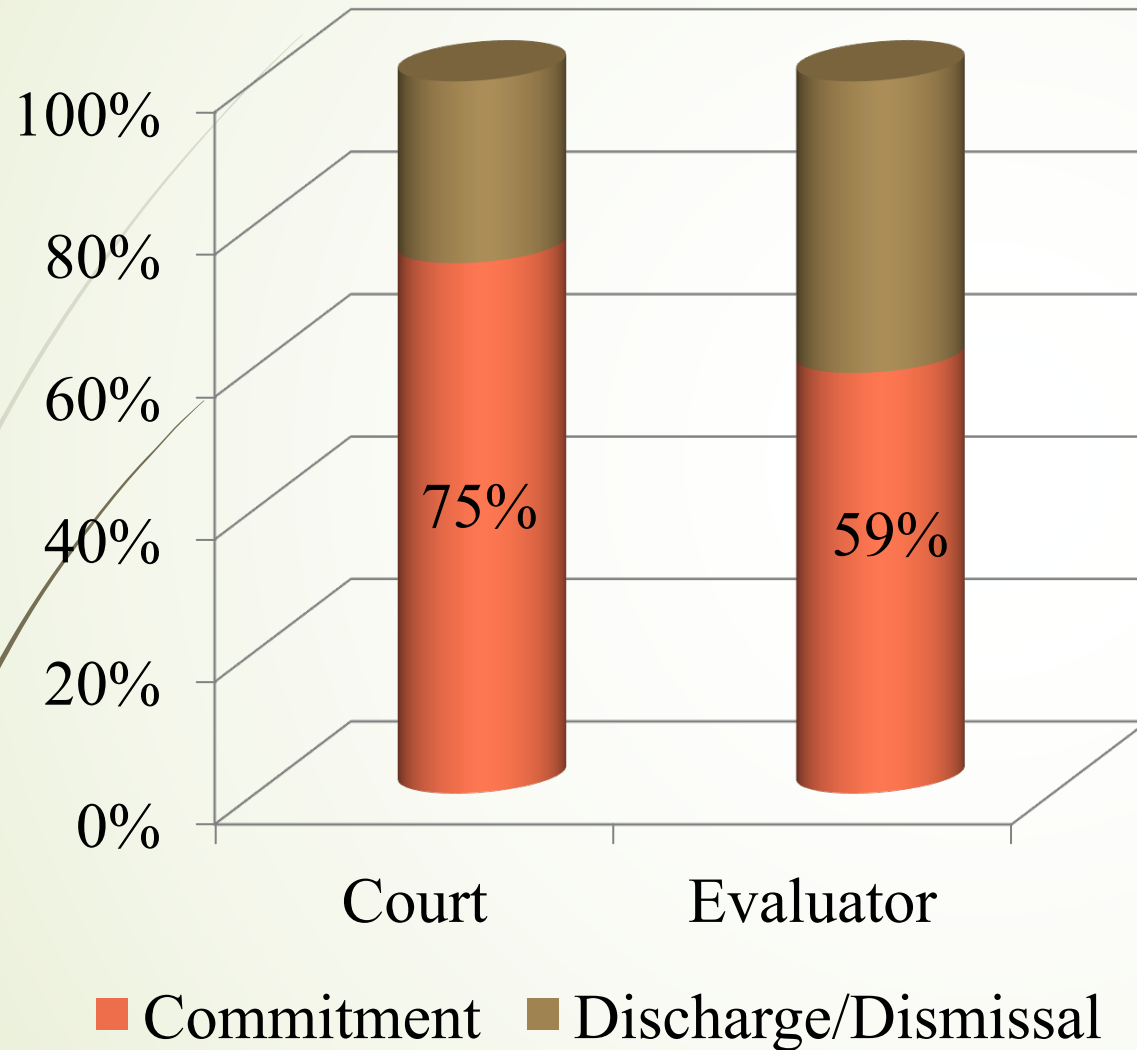


	N (%)
Case Type	
Pre-Commitment (04)	85 (45%)
Post-Commitment (07)	103 (55%)
Trial Type	
Jury Trial	70 (37%)
Bench Trial	118 (63%)
Stipulated Agreement	
Yes	56 (30%)
No	132 (70%)



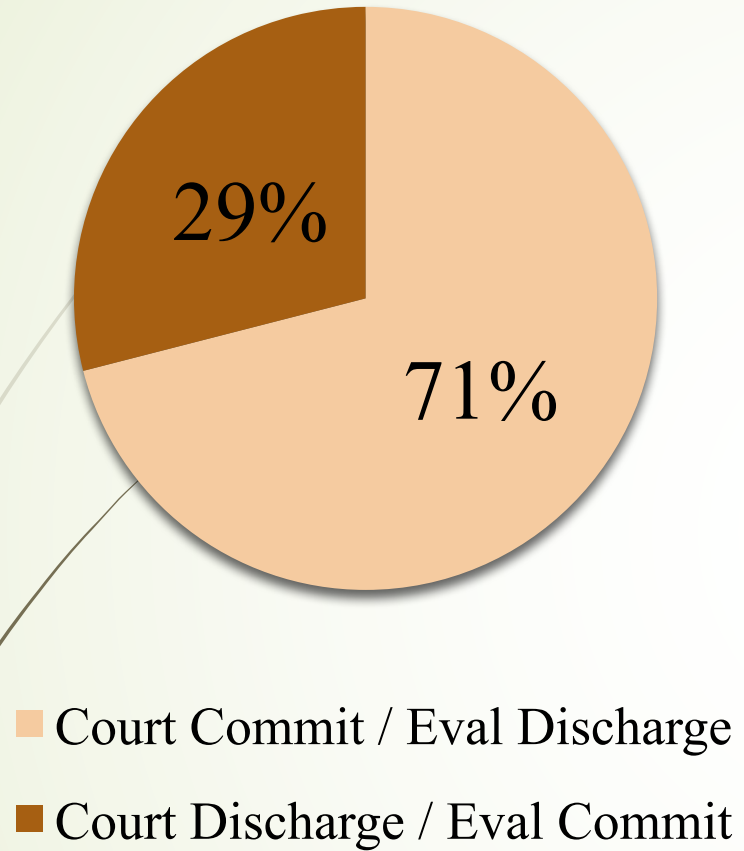
How often do courts agree with a DHS
state evaluator?

All Court Outcomes (n = 188)



■ Agreement ■ Disagreement

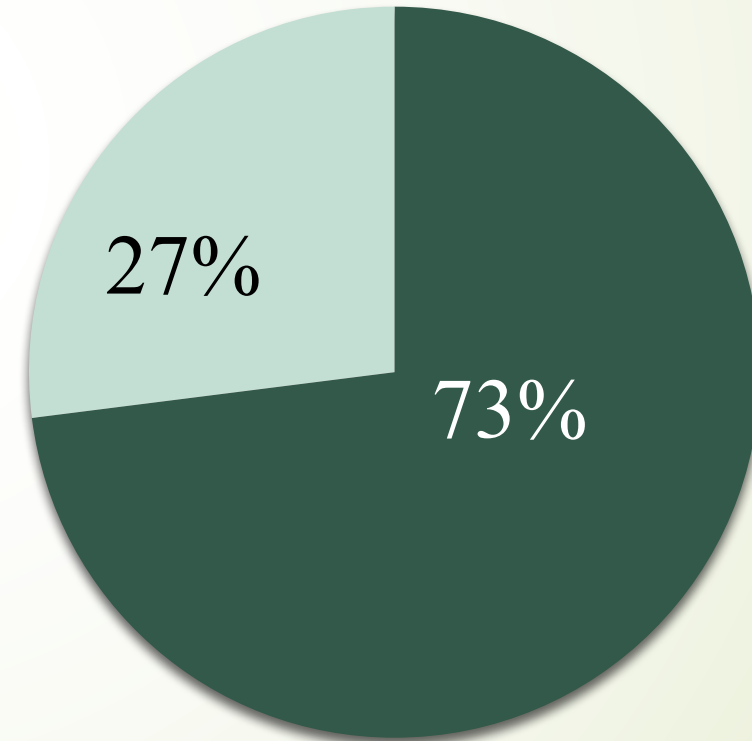
Disagreements (n = 51)



McNemar Test, $p < .001$

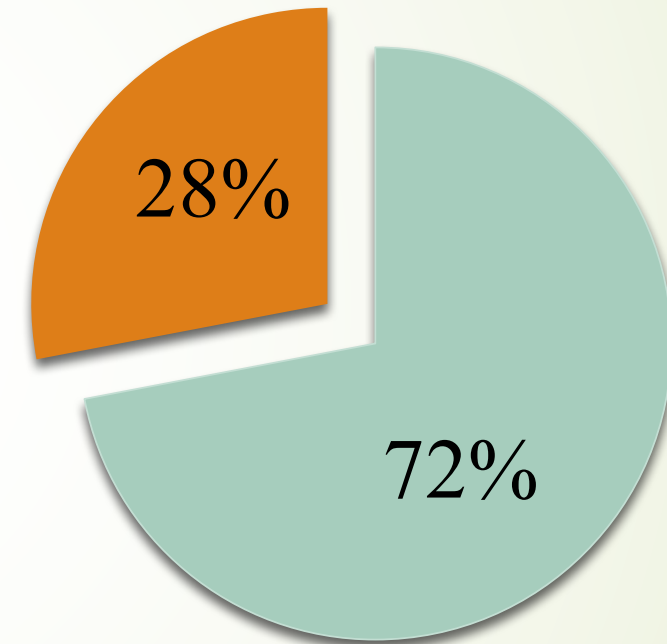
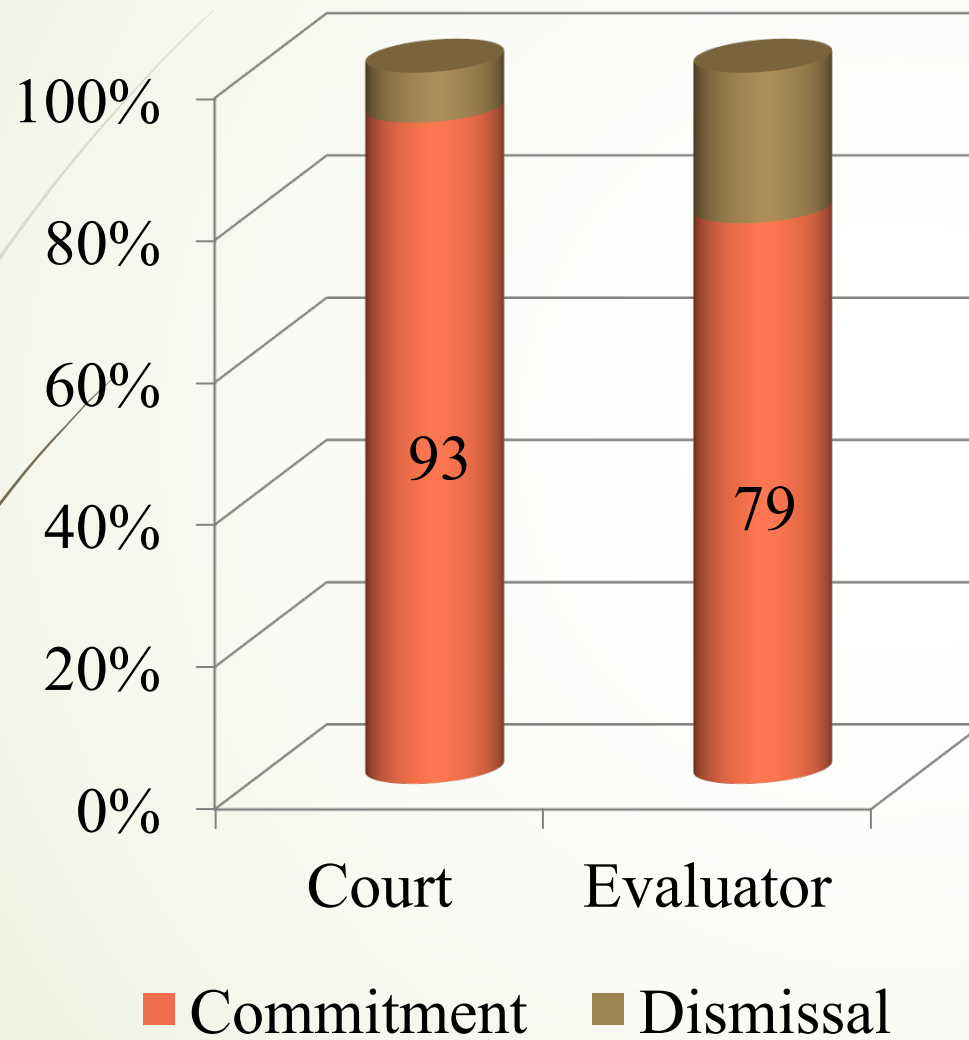
$K = .405, p < .001$

Agreements (n = 137)



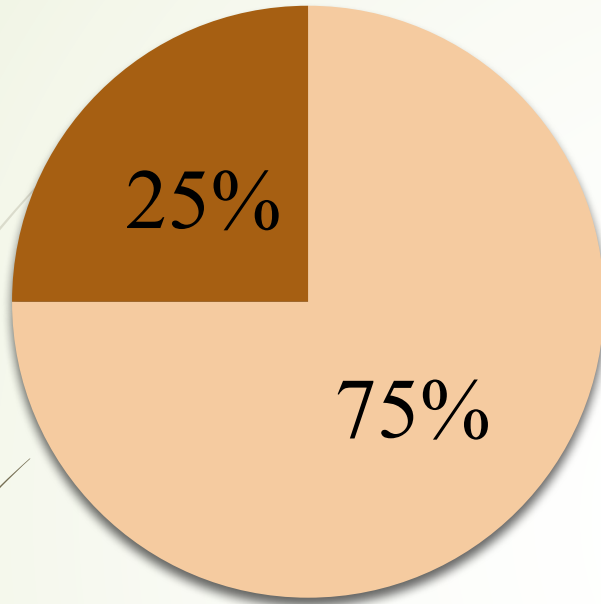
■ Agree to Commit ■ Agree to Discharge

Pre-Commitment Proceedings (n = 85)



■ Agreement ■ Disagreement

Disagreements (n = 24)



■ Court Commit / Eval Discharge

■ Court Discharge / Eval Commit

McNemar Test, $p = .02$

$K = -.118, p = \text{n.s.}$

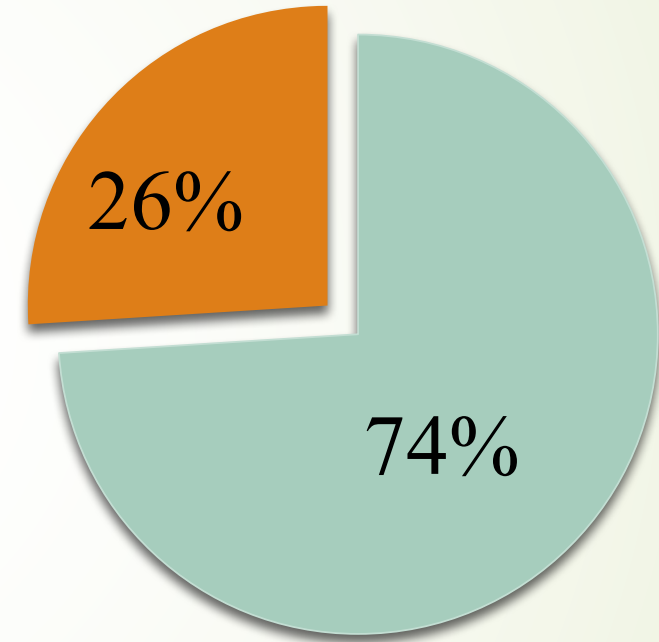
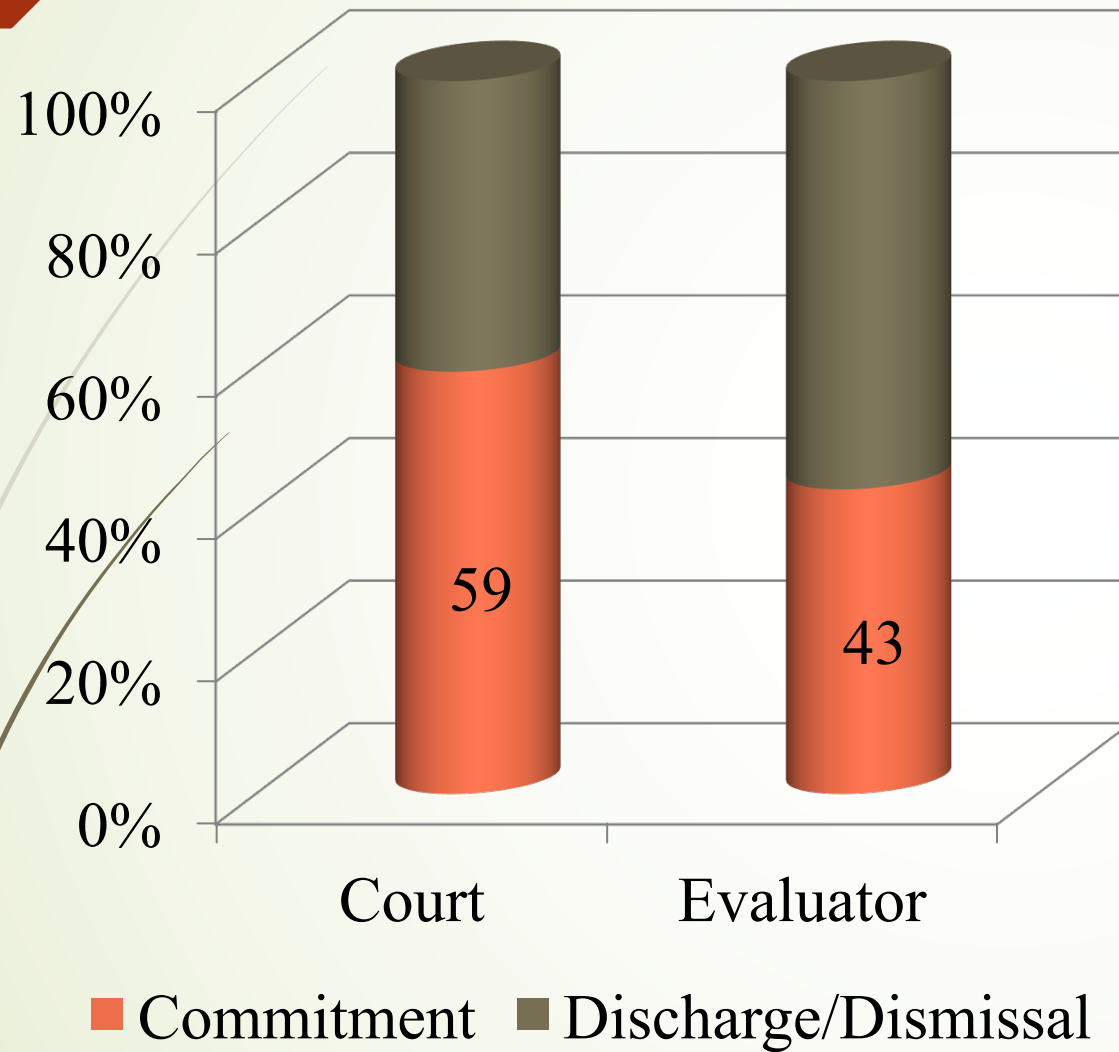
Agreements (n = 61)



■ Agree to Commit

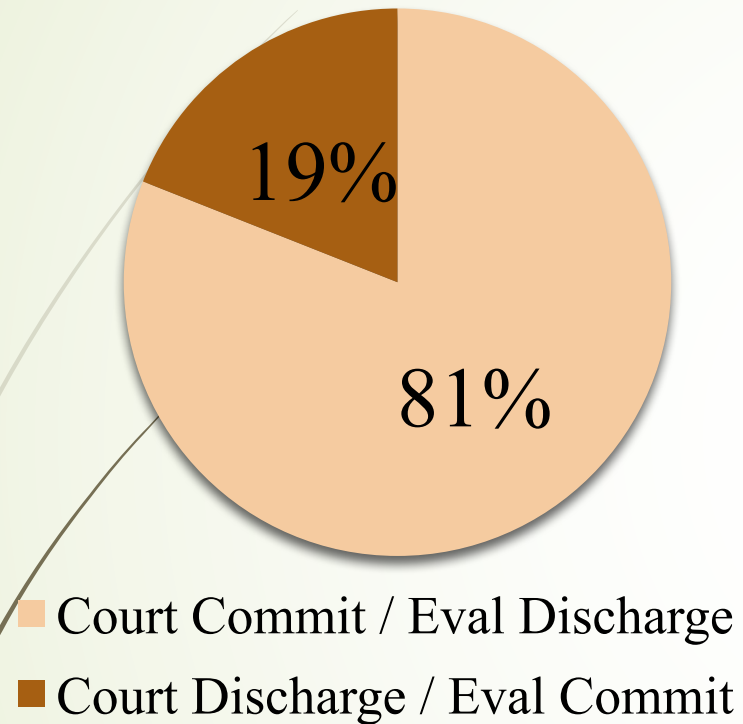
■ Agree to Discharge

Post-Commitment Proceedings (n = 103)



■ Agreement ■ Disagreement

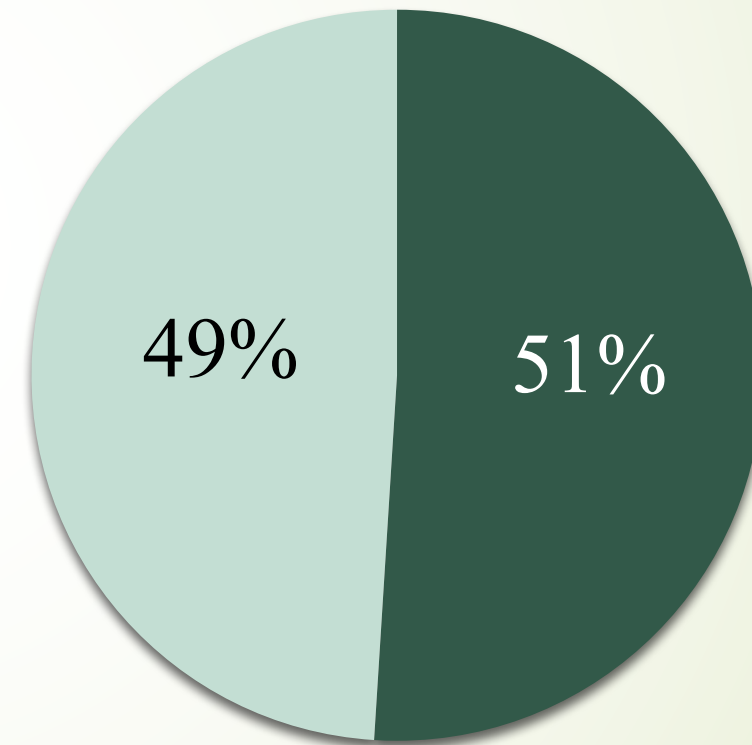
Disagreements (n = 27)



McNemar Test, $p = .002$

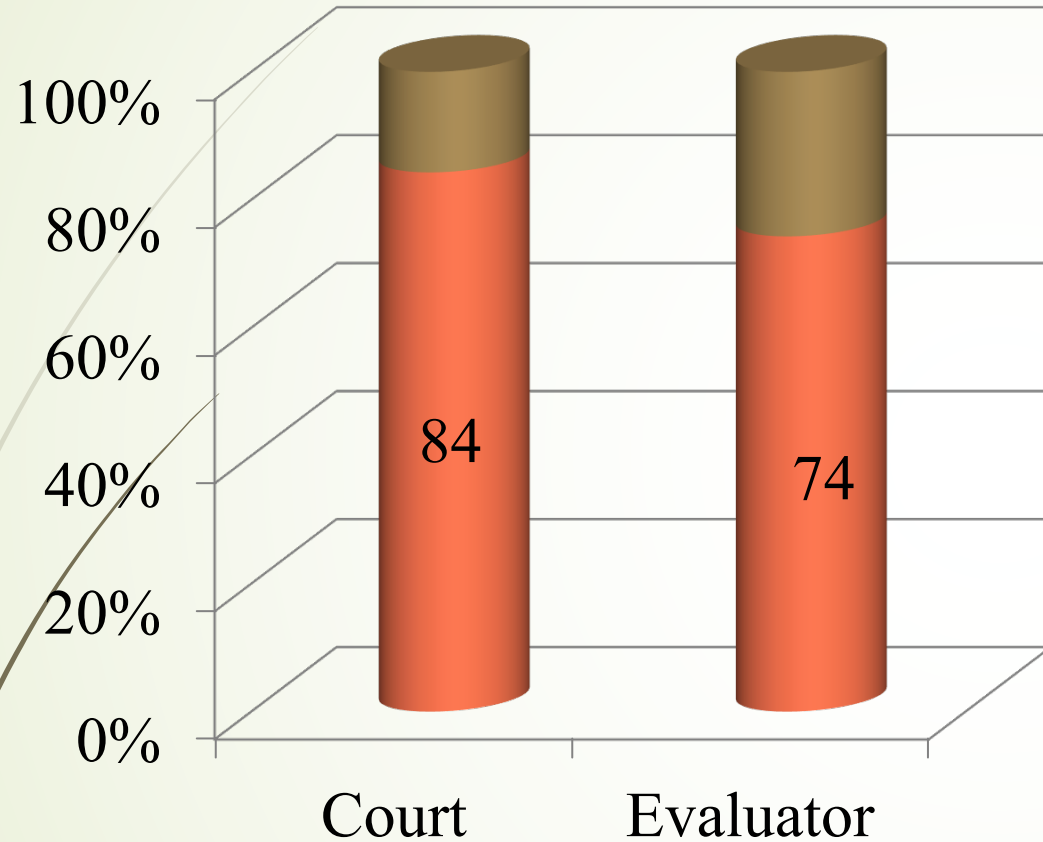
$K = .489, p < .001$

Agreements (n = 76)

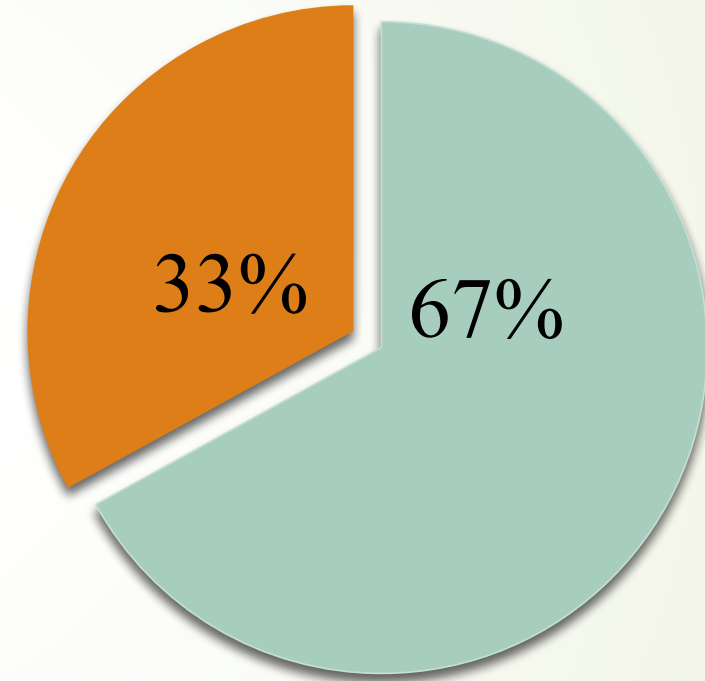


■ Agree to Commit ■ Agree to Discharge

Jury Decisions (n = 70)

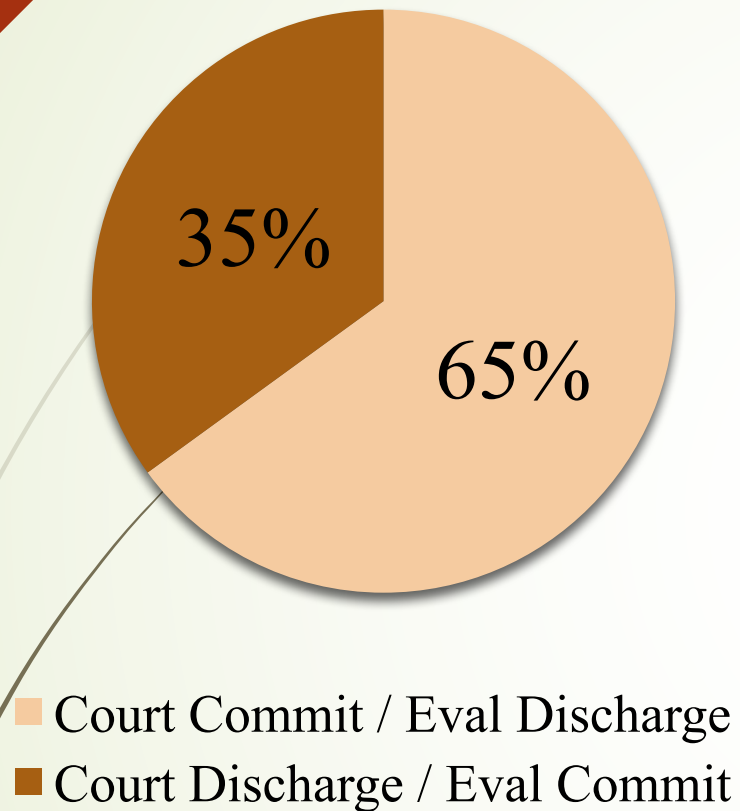


■ Discharge/Dismissal ■ Commitment



■ Agreement ■ Disagreement

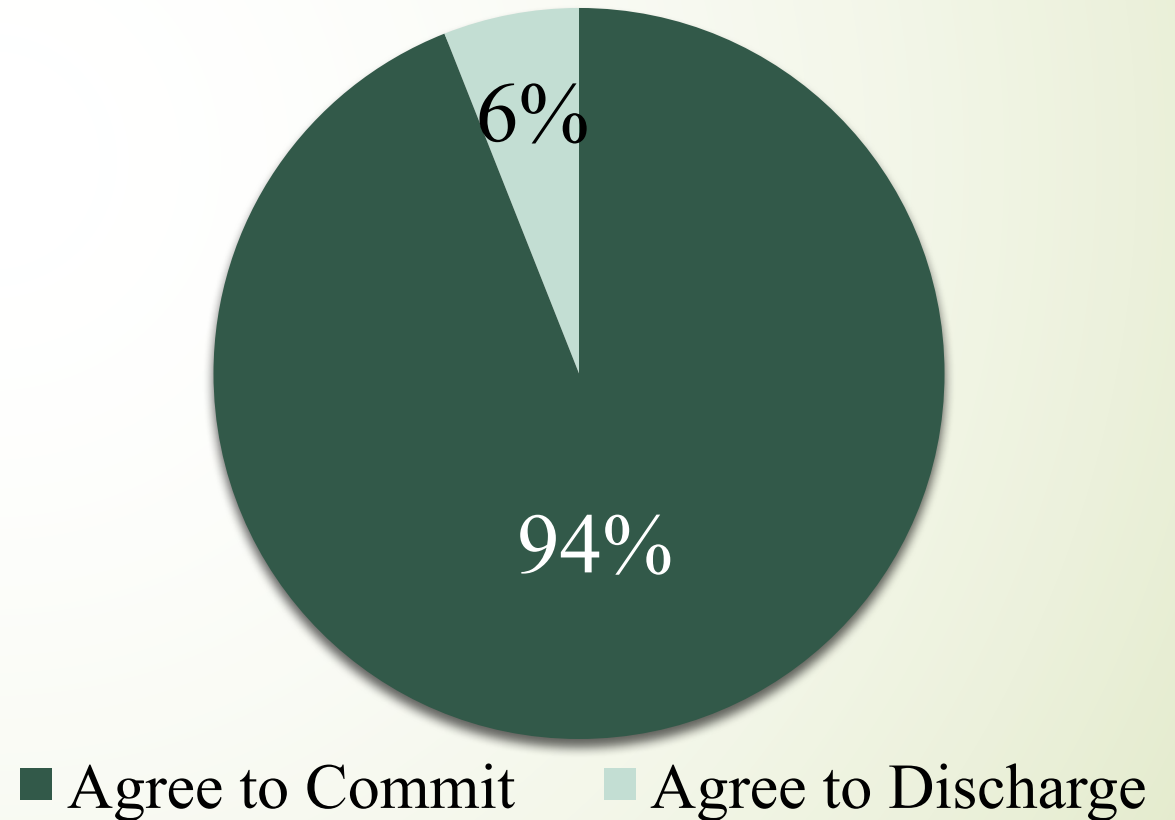
Disagreements (n = 23)



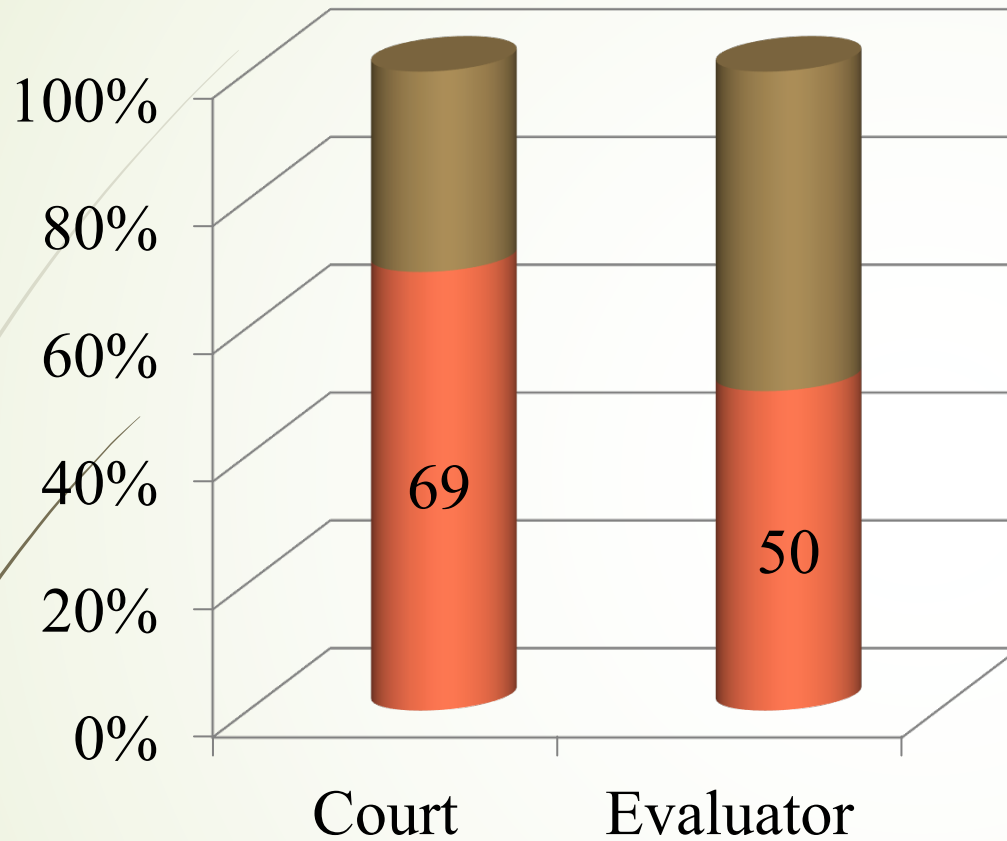
McNemar Test, $p = .21$

$K = .015, p = \text{n.s.}$

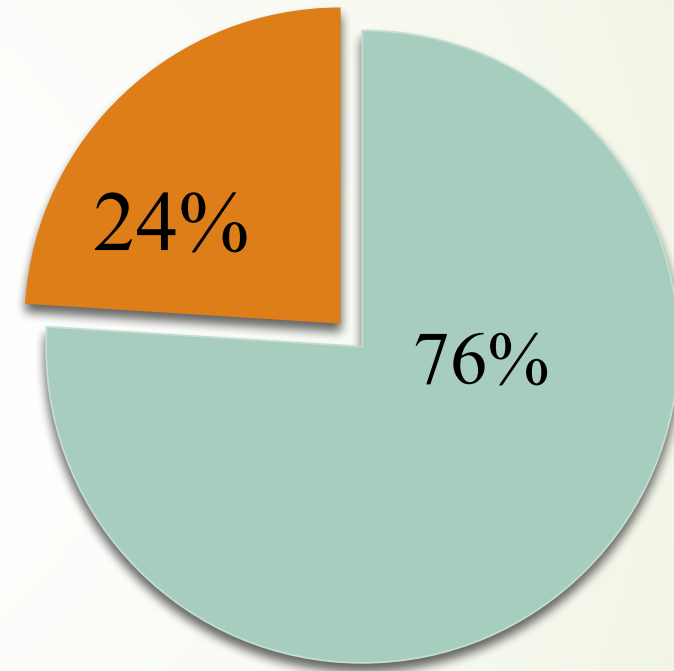
Agreements (n = 47)



Bench Decisions (n = 118)

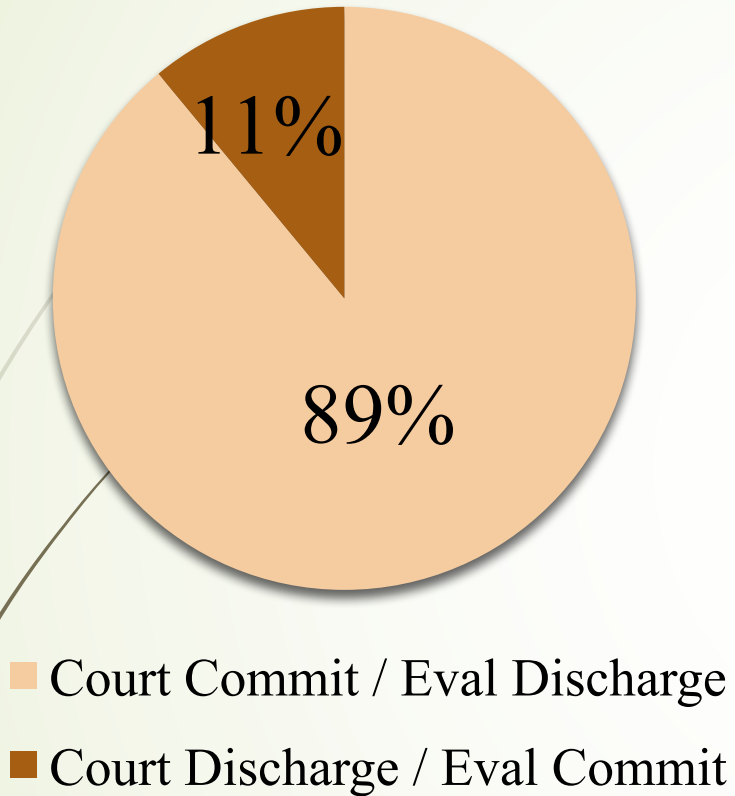


■ Discharge/Dismissal ■ Commitment



■ Agreement ■ Disagreement

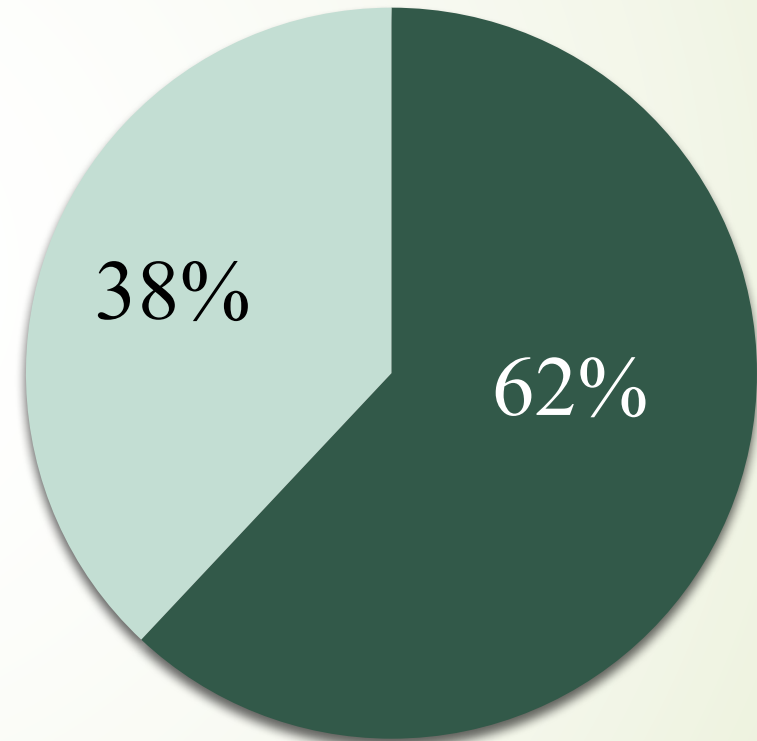
Disagreements (n = 28)



McNemar Test, $p < .001$

$K = .525, p < .001$

Agreements (n = 90)



■ Agree to Commit ■ Agree to Discharge



Summary Points on Agreement

- 77 of 188 (41%) patients were recommended for dismissal/discharge by evaluators
- 48 of 188 (26%) patients were dismissed/discharged by the court.
- Across all trials there is “moderate agreement” between courts and state evaluators
 - Agreement is poorer for pre-commitment cases and jury rulings
 - Substantially higher in post-commitment cases and bench rulings
- Disagreements were most frequently courts reaching a commitment decision despite DHS evaluators recommendation for discharge/dismissal.



Does evaluator recommendation predict
court outcome?

Does DHS evaluator recommendation predict the court decision? ($n = 188$)

	B (SE)	Lower	Odds Ratio	Upper
Jury / Bench	.14 (.46)	.46	1.15	2.85
Type of Case (Pre/Post)	1.72 (.51)*	2.06	5.56	15.03
Static99R	.01 (.14)	.76	.1.01	1.34
PCL-R Total	-.03 (.04)	.90	.97	1.05
Evaluator Recommendation	1.73* (.50)	2.12	5.65	15.07

For post-commitment trials alone? ($n = 103$)

	B (SE)	Lower	Odds Ratio	Upper
Jury / Bench	.48 (.62)	.48	1.62	5.44
Static99R	-.17 (.18)	.70	.98	1.38
PCL-R Total	-.16 (.05)	.90	.98	1.08
Evaluator Recommendation	2.50* (.65)	3.40	12.14	43.34

For Bench Decisions alone? ($n = 118$)

	B (SE)	Lower	Odds Ratio	Upper
Case Type (Pre/Post)	2.87* (1.08)	2.15	17.70	145.91
Static99R	-.03 (.19)	.68	.98	1.41
PCL-R Total	-.02 (.05)	.89	.98	1.08
Evaluator Recommendation	2.76* (.79)	3.39	15.82	73.83

For Jury Decisions alone? ($n = 70$)

	B (SE)	Lower	Odds Ratio	Upper
Case Type (Pre/Post)	.81 (.68)	.59	2.24	8.55
Static99R	.01 (.27)	.60	1.00	1.69
PCL-R Total	-.07 (.08)	.79	.93	1.08
Evaluator Recommendation	.14 (.86)	.22	1.15	6.16



Summary Points

- Evaluator recommendation and case type are significant predictors of ultimate court decision.....
 - After controlling for Static99R, PCL-R, and trial type
 - **Post-Commitment Bench Trials**
- No variables predicted jury trial outcome
- Risk assessment instruments were not predictive of court outcome after accounting for evaluator recommendation.

Is there anything unique about those patients who were dismissed/discharged ($n = 48$)?

➤ No differences in age, ethnicity, diagnoses

	Discharged/Dismissed ($n = 48$)		Committed ($n = 140$)			
	M	SD	M	SD	T-test	Cohen's d
Static99R	4.62	1.48	5.63	1.71	-3.64***	.63
Days since Admission	4151	1789	2426	2093	5.10***	.89
Days in Treatment	2450	1799	1340	1652	3.93***	.64



Summary

- Clearly commitment is a more likely outcome for patients under the 980 law- particularly at the pre-commitment stage
- Agreement between courts and state evaluators is variable
- State evaluator recommendation is a significant predictor for post-commitment trials – primarily bench trials
- Static99R is not a predictor of court outcome in logistic regression
 - Consistent with past research (*Boccaccini et al., 2013; Krauss et al., 2011; McCabe et al., 2010; Turner et al., 2015*)
 - Is a predictor of evaluation recommendation
 - Is significantly lower for discharged/dismissed patients



Main Limitations

- Cases sampled from 2012 to 2016 in Wisconsin
- Inability to capture other relevant variables
 - Other expert witnesses opinions
 - Attorney experience
 - Conservativeness of court
 - Perceived quality of testimony
 - Severity of crimes

Future Directions

- Collection of additional years of trial outcomes
- Evaluation of perceived expert witness credibility and efficacy
- Jury / Judge decision making questionnaires
- Comparison to other state rates



THANK YOU

► Co-Collaborators:

Richard Elwood, Ph.D.

James Mundt, Ph.D.

