The Impact of Sexual Assault Nurse Examiner Programs on Criminal Justice Case Outcomes: A Multisite Replication Study

Rebecca Campbell, Deborah Bybee, Stephanie M. Townsend, Jessica Shaw, Nidal Karim and Jenifer Markowitz

Violence Against Women 2014 20: 607 originally published online 28 May 2014
DOI: 10.1177/1077801214536286

Violence Against Women
2014, Vol. 20(5) 607–625

DOI: 10.1177/1077801214536286
vaw.sagepub.com
SAGE

The Impact of Sexual Assault Nurse Examiner Programs on Criminal Justice Case Outcomes: A Multisite Replication Study

Rebecca Campbell[1], Deborah Bybee[1], Stephanie M. Townsend[2], Jessica Shaw[1], Nidal Karim[3], and Jenifer Markowitz[4]

Abstract

To address the underreporting and underprosecution of adult sexual assaults, communities throughout the United States have implemented multidisciplinary interventions to improve postassault care for victims and the criminal justice system response. One such model is the Sexual Assault Nurse Examiner (SANE) Program, whereby specially trained nurses provide comprehensive psychological, medical, and forensic services for sexual assault. In this study, we conducted a multisite evaluation of six SANE programs (two rural programs, two serving midsized communities, two urban) to assess how implementation of SANE programs affects adult sexual assault prosecution rates. At each site, most sexual assaults reported to law enforcement were never referred by police to prosecutors or were not charged by the prosecutor's office (80%-89%). Individually, none of the sites had a statistically significant increase in prosecution rates pre-SANE to post-SANE. However, when the data were aggregated across sites, thereby increasing statistical power, there was a significant effect such that cases were more likely to be prosecuted post-SANE as compared with pre-SANE. These findings suggest that the SANE intervention model does have a positive impact on sexual assault case progression in the criminal justice system.

[1]Michigan State University, East Lansing, USA
[2]Townsend Consulting & Evaluation, Inc., Pittsford, NY, USA
[3]CARE USA, Atlanta, GA, USA
[4]Forensic Nurse Consultant, Alexandria, VA, USA

Corresponding Author:
Rebecca Campbell, Department of Psychology, Michigan State University, 127 C Psychology Building, East Lansing, MI 48824-1116, USA.
Email: rmc@msu.edu

Nevertheless, there is still a pressing need for improvement as the vast majority of both pre-SANE and post-SANE resulted in nonreferral/no charges filed.

Keywords
criminal justice case progressions, Sexual Assault Nurse Examiner (SANE), sexual assault prosecution

Sexual violence is a pervasive social problem: national epidemiological data indicate that 18% to 25% of women are raped or sexually assaulted in their adult lifetimes (Black et al., 2011; Kilpatrick, Resnick, Ruggiero, Conoscenti, & McCauley, 2007; Tjaden & Thoennes, 2006).[1] Despite the alarming prevalence of this crime, most sexual assault victims do not report to law enforcement (Planty, Langton, Krebs, Berzofsky, & Smiley-McDonald, 2013), and of those incidents that are reported, the vast majority will not be prosecuted (Campbell, 2008). In efforts to address this "justice gap" (Lonsway & Archambault, 2012; Spohn & Tellis, 2012a), many communities throughout the United States have implemented multidisciplinary interventions to improve postassault care for survivors and increase reporting and prosecution rates (Campbell, Patterson, & Bybee, 2012).

One such intervention model is Sexual Assault Nurse Examiner (SANE) programs, whereby specially trained nurses provide comprehensive psychological, medical, and forensic services for sexual assault victims in either hospital or community-based clinic settings (Department of Justice [DOJ], 2013). These programs first emerged in the 1970s and proliferated rapidly in the1990s and early 2000s, now numbering more than 700 SANE programs in the United States (International Association of Forensic Nurses [IAFN], 2013). Sexual assault forensic nurses are trained to offer crisis intervention and emotional support, health care (e.g., sexually transmitted infection [STI] screening and prophylaxis, pregnancy testing and emergency contraception), injury detection and treatment, and state-of-the-art forensic medical evidence collection. In addition, SANEs work with police and prosecutors to provide ongoing case consultation and may testify as expert witnesses should a case go to trial (DOJ, 2013).

Several case study evaluations of SANE programs conducted in the 1980s and 1990s suggested that these interventions can significantly increase arrest and prosecution rates (see Campbell, Patterson, & Lichty, 2005, for a review); however, these evaluations did not include comparison groups or other methodological controls, making it difficult to ascertain whether the reported increases were in fact attributable to the implementation of SANE programs. To date, there have been only two studies that have used rigorous pre–post quasi-experimental designs to evaluate whether SANE programs can significantly increase prosecution rates.[2] Crandall and Helitzer (2003) compared prosecution rates in an urban New Mexico jurisdiction 2 years before the implementation of a SANE program to 3 years after the program started and found significant increases in reporting, arrest, charging decisions, prosecutions,

and convictions. Similarly, Campbell et al. (2012) compared investigation and prosecution rates 5 years before the implementation of a Midwestern suburban SANE program to rates 7 years after the program was instituted. Multilevel modeling revealed that investigation and prosecution rates significantly increased post-SANE: more sexual assault cases were moving further through the system, reaching higher levels of case disposition (i.e., plea bargains and trials) after the implementation of the SANE program. Due to the methodological strength of the quasi-experimental design and supplemental data collection used in that project, these effects could be reasonably attributed to the efforts of the SANE program and not due to other county-level effects over time.

However promising these findings may be, they must be interpreted with caution because whether such gains are typical across different settings and contexts is not yet known. Research on the dissemination of innovative interventions suggests that positive outcomes may not be universal; program implementation often varies with respect to structure, function, and operation, and those differences are often critical in whether interventions achieve success (see Rogers, 2003). Indeed, in a national random survey of SANE programs, Patterson, Campbell, and Townsend (2006) found substantial variability in program practice philosophy, the nature of collaboration with the legal system, as well as other factors that may be critical for increasing prosecution rates (Campbell, Patterson, & Fehler-Cabral, 2010). As such, there is a pressing need for multisite replication studies to examine whether the positive results found in some jurisdictions are evident in other communities as well.

Therefore, the purpose of the current study was to conduct a national, multisite study to examine whether the implementation of SANE programs contributes to increased sexual assault prosecution rates. Based on the work by Campbell et al. (2012), we conducted a *conceptual replication*, whereby we retained the same focal research question, but deliberately changed both the study's settings and some elements of the methodology (Schmidt, 2009). Conceptual replications are an ideal choice when the goal is to determine whether an established effect will emerge under different conditions (as opposed to *operational replications*, whereby the methods are repeated as closely as possible to determine *if* the effect can be reproduced at all; see Schmidt, 2009, for a review). SANE programs operate in markedly different contexts with respect to geographic location, program setting, and number of patients served, but to date only large urban/suburban programs have been studied. As such, we wanted to vary program setting to examine the impact of SANE programs in rural, midsized, *and* urban communities.

Furthermore, given the number of SANE programs in existence (and their continued rapid proliferation), a traditional research design whereby one site/one program is studied in-depth may not be an efficient strategy for addressing the larger question of whether these programs, as an intervention model, are effective. Wandersman (2003) and Wandersman et al. (2008) noted that when a new service model "takes off" as quickly as the SANE approach has, it may be more fruitful to focus on building evaluation capacity at the local level and to engage program stakeholders to co-conduct program evaluations. By engaging practitioners in the process of evaluating their

work, the knowledge base can develop much more rapidly—and the findings are more likely to be used to improve practice, as local stakeholders are directly involved in the project (see Cousins & Chouinard, 2012; Patton, 2008, for reviews). Consistent with that approach, our conceptual replication involved building partnerships with multiple SANE programs to work with program staff to co-collect data regarding sexual assault prosecution rates in each jurisdiction.

Bringing these ideas together, the purpose of this project was to conduct participatory evaluations with six SANE programs: two rural programs, two programs serving midsized communities, and two urban programs. We provided a comprehensive training and technical assistance package to program staff, which included a step-by-step evaluation toolkit, multiple training webinars, group consultation calls, individualized phone and email consultation, and in-person site visits (see Campbell, Townsend, Bybee, Shaw, & Markowitz, 2013 for more details). Using standardized sampling and data collection methods comparable with Campbell et al. (2012), program staff collected criminal justice case outcomes for sexual assault cases in which victims had a medical forensic exam and reported to police, comparing case outcomes before and after the implementation of the SANE program in their communities. In this article, we will highlight the results from the site-specific and cross-site compiled data regarding the impact of SANE programs on sexual assault prosecutions rates.

Method

Site Selection

We used stratified random sampling to identify six SANE programs (two rural, two midsized, two urban) that had organizational readiness to participate in program evaluation. Consistent with Preskill's model of evaluation capacity building (Preskill & Boyle, 2008; Russ-Eft & Preskill, 2001), we distinguished evaluation "readiness" from evaluation "capacity." Briefly, "readiness" means that the program had the organizational resources to participate in evaluation learning activities without compromising program operations; "capacity" means program staff have already learned evaluation skills and routinized evaluation within their organization. Sampling programs based on evaluation *readiness* was appropriate because we wanted to target more typical program conditions in which the program had the organizational resources to support evaluation efforts but did not necessarily have the experience or expertise in conducting evaluation.

We chose to sample based on evaluation readiness, rather than draw a random sample from all U.S. SANE programs because we knew that participating in this project would take staff time and effort and we did not want the evaluation to jeopardize patient care. Prior evaluation projects have documented that many SANE programs are facing sustainability challenges and are struggling to provide consistent 24/7 high-quality care for their patients (Campbell & Patterson, 2010); as such, it seemed likely that a national random sample could have selected a program (or programs) that would not have been able to complete the evaluation. Given that our goal was to explore how

to build evaluation capacity in SANE programs, the sampled programs had to have the organizational readiness for such work. Therefore, our sampling strategy focused on identifying programs that had the potential to complete an evaluation and then we randomly selected among those programs, stratifying by community/program size to ensure that we had representation from rural, midsized, and urban programs.

To recruit programs and screen for organizational readiness, we collaborated with the IAFN and the National Sexual Violence Resource Center (NSVRC) to advertise this project to all SANE programs within the United States, which at the time this study was conducted numbered approximately 600 programs. Interested SANE programs were instructed to contact the research team for a readiness assessment application. Seventy-three SANE programs expressed interest in becoming a study site: 30 programs completed the entire online application process; 7 programs had incomplete applications; and 36 programs did not activate their online application (i.e., they elected not to apply after expressing initial interest).[3] The 30 complete applications were independently reviewed by each member of the research team (which included a highly experienced forensic nurse consultant) to assess eligibility. Programs were deemed eligible to participate if they had (a) a full-time SANE program coordinator; (b) nursing staffing levels appropriate for their current patient caseloads so that staff would be able to devote time to participating in an evaluation project without adversely affecting program services; (c) a mean score above the 25th percentile on selected items from Preskill and Torres' (2000) Readiness for Organizational Learning and Evaluation Instrument (ROLE) scale, which indicates good organizational readiness for evaluation activities; and (d) secured access to the different data sources needed to complete a pre-SANE/post-SANE evaluation of legal case outcomes. Of the 10 eligible programs, a stratified random sample was drawn consisting of two rural sites (Sites A and B), two midsized sites (Sites C and D), and two urban sites (Sites E and F).

After site selection, we discovered that three sites would not be able to conduct a pre-SANE/post-SANE evaluation design due to unforeseen challenges in accessing pre-SANE data.[4] Therefore, these three sites needed to do a post-only evaluation design. These sites were evenly distributed across the sampling strata (i.e., one rural, one midsized, and one urban SANE program completed post-only designs) and our analysis plan was adjusted accordingly to accommodate this design change (see below). Each SANE program obtained institutional review board (IRB) approval from its parent hospitals or host organizations given that they would be accessing patients' medical records.[5] In addition, each site established a memorandum of understanding (MOU) with its local prosecutor's office to access case outcome records. Sites did not receive financial compensation for their participation (only technical assistance).

Site Training and Technical Assistance

To ensure fidelity to the research design and data collection procedures, all sites participated in a series of web-based trainings following their enrollment in the project. The trainings included live audio presentation, visuals, and discussion transmitted via an online conferencing platform. The trainings consisted of a series of three hour-long

sessions: (a) an overview of the project and research process, (b) an explanation of the first steps of the research design (understanding the design, identifying evaluation questions, establishing cooperative agreements, sampling cases, and collecting data), (c) an explanation of the last steps of the research design (analyzing data and interpreting results). In addition, all-site conference calls were held periodically throughout the project to check on progress, share successes, and troubleshoot challenges. These calls were augmented by individual communication (via phone and email) between sites and the research team (approximately 450 email threads throughout the project). In addition, in-person site visits were conducted at each program to review the data that had been collected, assist with data analysis and interpretation, and develop action plans based on the findings.

Sampling Design and Case Selection

After completion of the training webinar on sampling, program staff were instructed to review their files to identify eligible cases for selection. Following the same sampling procedures established by Campbell et al. (2012), pre-SANE and post-SANE cases needed to meet the following criteria to be eligible for study inclusion: (a) the patient was at least 18 years old at the time of the sexual assault; (b) the patient had a complete medical forensic exam and evidence collection; (c) the exam and evidence collection were not anonymous/de-identified; (d) the patient made a police report; (e) the police report was not anonymous or "Jane Doe"; and (f) the assault occurred within the focal county (counties) that were the subject of the evaluation.

Each site also had to determine the time period from which cases would be selected. Sites completing a post-only evaluation design selected cases that received services from their SANE program beginning 1 year after the program start date through 1 year prior to the evaluation (i.e., the first year and most recent year of SANE program operation were excluded).[6] Sites completing a pre–post evaluation design selected the same post-SANE time period; the pre-SANE period was defined as 3 (or 2, depending on record availability) years prior to the launch of their program.

As collecting prosecution case outcomes is a resource-intensive process (see Campbell, Bybee, Ford, Patterson, & Ferrell, 2009; Spohn & Tellis, 2012b), it was necessary to limit the number of cases studied, particularly in the context of program evaluations conducted by practitioners. We conducted a power analysis to determine the minimum number of cases per year that could be sampled and still yield .80 power for within-site and cross-site analyses. This analysis revealed that 30 cases per year would be necessary, which was also a programmatically reasonable number for program staff. Sites that had more than 50 eligible cases per year randomly sampled 30 cases per year for inclusion in this project. The research team provided intensive training to each site that included standardized directions and step-by-step forms to complete when reviewing case files to determine case eligibility and sample cases. The research team monitored sampling throughout the project via group conference calls and individual technical assistance; we continually reviewed the sampling criteria to make sure staff understood the methods and we invited discussion regarding any

Table 1. Summary of Research Designs and Samples for the Six Selected SANE Programs.

	Design	Number of years pre	Number of years post	Sample size pre	Sample size post
Site A (rural)	Pre/post	3	5	47	92
Site B (rural)	Post-only	NA	9	NA	253
Site C (midsized)	Pre/post	3	10	65	334
Site D (midsized)	Post-only	NA	7	NA	201[a]
Site E (urban)	Pre/post	2	10	60[a]	300[a]
Site F (urban)	Post-only	NA	12	NA	344[a]

Note. SANE = Sexual Assault Nurse Examiner.
[a]Sampled 30 cases per calendar year.

unusual cases that might have come up. We developed (and continually revised) a frequently asked questions (FAQ)-type document that highlighted common (and unusual) sampling issues. Table 1 provides a summary of the evaluation designs and sampling/sample sizes in each of the six sites.

Data Collection Procedures and Measures

All sites were provided with standardized instructions and data collection forms to obtain case outcome data from the prosecutor's office for all selected cases. Consistent with measures used by Campbell et al. (2012), case outcomes were classified into one of five mutually exclusive categories: (a) the case was not referred by police/not charged by prosecutors; (b) the case was charged by prosecutors, but later dropped; (c) the case was plea bargained; (d) the case went to trial and ended in an acquittal; or (e) the case went to trial and ended in a conviction.[7] Unknown or pending case outcomes were also recorded, although there were few instances of that as sampling was designed to minimize missing data. Cases still pending or with otherwise unknown outcomes were excluded from the analyses.[8] As noted previously, the research team monitored data collection very closely via group conference calls and individual technical assistance to ensure that program staff were coding case outcomes consistently.

Results

Table 2 summarizes the pre- and post-SANE prosecution outcomes for each of the six sites. Despite marked differences in program geography/setting, all six programs had remarkably similar findings. In the three sites that were able to complete a pre–post evaluation design, the pre-SANE rates of nonreferral/not charged were strikingly consistent: 89%, 94%, and 87% of cases were not referred/not charged. In other words, in cases of adult sexual assault in which the victim had obtained a medical forensic exam and had reported the assault to the police, it was extremely unlikely that legal action was taken (90% on average, across sites). After the implementation of the SANE programs in these three communities, the rates of nonreferral/not

Table 2. Legal Case Outcome Findings, By Site.

	Site A (rural)		Site B (rural)	Site C (midsized)		Site D (midsized)	Site E (urban)		Site F (urban)
	Pre (%)	Post (%)	Post (%)	Pre (%)	Post (%)	Post (%)	Pre (%)	Post (%)	Post (%)
Not referred/ not charged	89	80	91	94	89	89	87	84	82
Charged, but later dropped	0	2	0	2	0	5	8	7	4
Plea bargained	11	13	4	3	6	4	3	7	13
Trial: acquittal	0	3	0	0	2	0	0	1	1
Trial: conviction	0	1	5	2	3	1	2	1	1

charged dropped (80%, 89%, 84%, respectively), indicating that more cases were progressing into the criminal justice system. However, the rates of nonreferral were still high (86%, on average across sites) and the change was not enough to reach statistical significance in any individual site. The rates for plea bargains and convictions generally increased post-SANE but again did not reach statistical significance in any individual site.

The three sites that conducted post-only evaluations had very similar results: 91%, 89%, and 82% of post-SANE cases were not referred/not charged (86% on average, across sites). The rate of trial convictions and plea bargains was also quite low (and again, consistent with the rates in the pre–post designs), with the notable exception that Site F (urban) had a markedly higher rate of plea bargains (13%).

Each of the six sites used identical data collection and coding procedures, so the data can be compiled for cross-site analyses to examine the impact of SANE programs on adult sexual assault prosecution rates.[9] The dependent variable—extent of progression through the system—was recoded to comprise three ordered categories characterizing the ultimate disposition of each case: (a) not referred/not charged (1,465 cases, 86.4%), (b) charged by the prosecutor but later withdrawn or acquitted (71 cases; 4.2%), and (c) guilty plea or conviction (160 cases; 9.4%). Cases spanned more than 14 years (169 months), from June 1995 through September 2009. Because cases handled near the same time may be influenced by secular trends and shared historical circumstances, cases were grouped for analysis by the calendar month in which examinations were conducted. The extent of variance in the dependent variable that was shared within month was modest, with the intraclass correlation coefficient (ICC) = .02, indicating that 2% of the variance among case outcomes could be explained by month-to-month fluctuations in case progression.

Grouping cases by month allowed examination of the data for time trends and possible seasonal effects that should be reflected in the analysis. Graphical inspection showed a slight but nonsignificant upward trend in the extent of progression toward prosecution across all 169 months. There was also a noticeable dip in level of system progression for exams conducted in the month of July. Across years and sites, July cases reached lower levels of progression through the system compared with cases presenting during the other months of the year (Somer's $d = -.06$, $p < .01$). Compared

with other months, cases processed in July were less likely to be charged (7.8% % vs. 14.1%) and less likely to result in a conviction or plea bargain (3.5% vs. 10.0%).

To reflect both the grouping of cases by month and the ordinal nature of the dependent variable, multilevel ordinal regression (Hedeker & Gibbons, 1994) was used to analyze the impact of SANE implementation on case progression through the justice system. Ordinal regression analyzes the cumulative probability that a case will exceed a particular level on the ordinal outcome variable, as a function of the case characteristics on explanatory variables included in the analysis. Multilevel ordinal regression is an extension that incorporates the shared influence of explanatory variables that affect groups of cases (in this research, cases handled in the same month), along with variables that exert independent effects on individual cases. Multilevel analyses produce standard errors that are appropriate for testing the influence of group-level variables, reflecting the lack of independence of their effects on individual cases. The current analysis reflected two levels of data: individual cases (Level 1), which were nested within months (Level 2). The 1,696 individual cases at Level 1 were nested within 169 months at Level 2. Modeling time at a higher level of analysis is somewhat unusual in multilevel analysis, but this strategy is described in Goldstein (2003) as appropriate for analyzing traditional time series data in which cases are nested within units of time.

One explanatory variable was examined at Level 2. This variable characterized the "July effect"—the dip in level of system progression for cases with examinations in the month of July (coded 1 for July month in each year; 0 for the other 11 months). Five explanatory variables were examined at Level 1. The first identified the site for each individual case; the six sites were characterized by five dummy variables. This variable reflected anticipated differences among the six sites in their handling and disposition of cases, and it also adjusted for confounding between time and site effects due to differences in month-to-month fluctuation of case distributions and site differences in the timing of SANE implementation. The second Level 1 variable differentiated the three sites with post-SANE only designs (Sites B, D, and F) from the three sites with pre-SANE/post-SANE designs (Sites A, C, and E). This variable was examined to determine if sites with different data collection designs differed significantly in overall level of case progression. It was not significantly related to case outcome, either as a main effect or in interaction with time, indicating no significant confounding between site design and case progression; therefore, this variable was omitted from further analysis.

The remaining Level 1 variables described the timing of each case. The third variable characterized the month in which each examination was conducted. To facilitate interpretation, months were numbered sequentially, centered so that 0 reflected February of 2003, the month of the initial SANE case for the last implemented SANE program. Linear, quadratic, and cubic effects of time were tested, in an effort to identify any underlying time trends in the progression of cases through the justice system. No effects approaching significance were found, either unconditional or conditional on the effects of other explanatory variables. Time-by-site interactions were also examined, to test for possible site differences in time trajectories; none was

significant. Because no significant effects of the continuous time trajectory were identified, this variable was omitted from further analysis.

The fourth variable examined the year of SANE program operation in which each examination was performed; values ranged from 0 (designating pre-SANE implementation years) to Year 13 (designating the 13th year of SANE program operation at a given site). Linear, quadratic, and cubic effects of this variable were examined to test for possible changes in case progression that could be attributable to program age. Program year-by-site interactions were also tested. No effects were found, and this variable was omitted from further analysis.

The fifth and final Level 1 variable identified months as pre-SANE (coded 0) versus post-SANE implementation (coded 1) and tested the central question of the analysis, whether there was a significant change in case progression following the implementation of SANE programs across sites. Site differences were also examined by testing site-by-SANE interactions for the three sites with both pre- and post-SANE implementation data.

Analyses were conducted with HLM 7 software (Raudenbush, Bryk, Cheong, & Congdon, 2011), using the hierarchical generalized linear model with a logit link function to characterize an ordinal dependent variable. Restricted maximum likelihood was used for estimation. To reduce the influence of nonnormal distributions, robust standard errors were used to compute confidence intervals. Models included random intercepts and random effects of SANE implementation.

Results of the multilevel ordinal regression are summarized in Table 3. The first block describes the ordinal effect—the expected probabilities of justice system dispositions at the two thresholds of the ordinal dependent variable, adjusting for the influence of other explanatory variables. The log odds of conviction or guilty plea versus other dispositions (i.e., acquitted/withdrawn or not referred/charged) was −3.010, which translates to an odds ratio (OR) of 0.049. This indicates that a pre-SANE case from Site C (the omitted comparison site), processed in a month other than July (i.e., a case with scores of 0 on all explanatory variables), was only 5% as likely (OR = 0.049) to attain a conviction/plea relative to other dispositions (i.e., acquittal/withdrawn or not charged). The threshold difference in the log odds of being charged versus attaining a conviction/plea was 0.422. The cumulative log odds that a case would be charged, regardless of ultimate outcome, is the sum of the log odds of the two thresholds—conviction/plea and the case being charged (−3.010 + 0.422), or −2.588, which translates to an OR = 0.075. This indicates that a case with scores of 0 on all explanatory variables (e.g., pre-SANE, Site C, non-July) was 7.5% as likely to be charged as not charged.

The second block contains the single Level 2 effect on month-to-month variation in case progression, the adjustment for the seasonal dip in case progression seen in the months of July. Across years and sites, cases with exams in July were less than half as likely (OR = 0.462; $p = .047$) to progress further in the system, compared with cases with exams in other months.

The third block of Table 3 shows the Level 1 (individual case) effects on the cumulative probabilities of justice system dispositions described above. The effects of site

Table 3. Cross-Site Results: The Impact of SANE Programs on Criminal Justice System Case Progression.

Fixed/average effects	Log odds	Robust *SE*	Odds ratio	*T*	*df*	*p*	Robust *CI*
Ordinal effect							
Threshold 1: Convicted/plead vs. not convicted (withdrawn/acquitted or not referred/charged)	−3.010	0.290	0.049	−10.392	168	.001	[0.028, 0.087]
Threshold 2: Charged (convicted/plead or withdrawn/acquitted) vs. not charged	0.422	0.051	1.525	8.316	1,356	.001	[1.380, 1.684]
Level 2 Effect							
Seasonality (July vs. other months)	−0.773	0.387	0.462	−2.000	167	.047	[0.215, 0.991]
Level 1 effects							
Site A vs. Site C	0.602	0.290	1.826	2.073	1,356	.038	[1.033, 3.229]
Site B vs. Site C	−0.230	0.261	0.795	−0.882	1,356	.378	[0.476, 1.325]
Site D vs. Site C	−0.010	0.271	0.990	−0.037	1,356	.970	[0.582, 1.684]
Site E vs. Site C	0.418	0.228	1.518	1.834	1,356	.067	[0.971, 2.373]
Site F vs. Site C	0.558	0.210	1.748	2.654	1,356	.008	[1.157, 2.641]
Post-SANE implementation vs. pre-SANE	0.584	0.271	1.793	2.148	167	.033	[1.048, 3.066]
Random effects variance estimates							
Intercept	0.312						
Post-SANE implementation vs. Pre-SANE	0.166						

Note. SANE = Sexual Assault Nurse Examiner; CI = confidence interval.

were assessed by five dummy variables. The site used as the omitted comparison (Site C) had a relatively low level of progression through the system, not significantly different from Site B and Site D. Cases at other sites were more likely to progress to higher levels: Site F (75% more likely; $p = .008$), Site A (83% more likely; $p = .038$), and Site E (52% somewhat more likely; $p = .067$).

The remaining effect at Level 1 is the comparison of post- versus pre-SANE implementation cases. After adjusting for other effects (seasonal variation and site), cases processed post-SANE implementation were almost 80% more likely (OR = 1.793) to attain a higher level of disposition, compared with cases processed pre-SANE; this effect was significant at $p = .033$. This effect was not found to differ significantly by site. In other words, although site-specific pre–post analyses were nonsignificant, when the data were compiled across sites to increase statistical power, there was a

significant effect such that adult sexual assault cases were significantly more likely to progress further through the criminal justice system after the implementation of a SANE program. This significant finding emerged even with the varying age of programs (i.e., the number of years data were collected per site) and number of case outcomes collected per site.

Discussion

Prior research suggests that SANE programs may be an effective intervention model for addressing the long-standing problem of sexual assault underprosecution. However, given that such conclusions are based on only a handful of studies—most of which have not used rigorous designs—there is a pressing need for multisite replication studies. In this project, we worked with a geographically diverse sample of SANE programs to examine how their services have affected prosecution rates in their local communities. Three of the six sites in this project conducted pre–post evaluations to ascertain whether adult sexual assault prosecution rates increased in their communities after the implementation of their SANE programs. None of the individual sites had significant findings, but when the data were aggregated across all sites, thereby increasing power, we found a statistically significant increase in case progression post-SANE: more cases were moving further through the system, reaching higher levels of case disposition (i.e., successful prosecution).

These positive multivariate findings regarding increased prosecution over time are consistent with the results of the Campbell et al. (2012) project; however, our site-level descriptive data tell a different—and far less positive—story. Pre-SANE prosecution rates in these three communities were quite low, and even though rates significantly increased, post-SANE rates were still rather dismal. For example, the vast majority of pre-SANE cases (90% on average) were *not* referred by the police to prosecutors or were not charged by prosecutors; post-SANE rates of nonreferral/noncharging significantly dropped to 86% (on average). The three sites that completed post-only evaluation designs also documented high rates of nonreferral/noncharging (also 86%, on average). Given that 100% of these cases were reported to police and 100% had a complete medical forensic exam and evidence collection, the fact that only 14% were even considered for prosecution is concerning. These descriptive data are markedly different than the rates documented in the urban/suburban community that was the focus of the Campbell et al. (2012) project. In that study, 34% of reported sexual assault cases with medical forensic exams were charged before the SANE program was implemented and that number increased to 42% (on average) post-SANE.

Taken together, the results of the original Campbell et al. (2012) study and this replication study suggest that SANE programs can help communities *improve* their prosecution rates for sexual assault, but there may be considerable variability in their baseline rates. In this multisite replication project, we sampled for greater geographic diversity (rural, midsized, and urban) and though we cannot pinpoint what specific aspects of these programs and their implementation varied from the interventions studied in prior research, we can reasonably conclude that there is greater heterogeneity in

SANE programs' successes than originally documented in the literature. One option for future research is to randomly sample from all U.S. SANE programs (stratified by geographic size/community context) to capture a nationally representative picture of their effectiveness. Such an approach has obvious methodological advantages, but given that many SANE programs have limited staffing (IAFN, 2013), time and effort invested in research/evaluation must be considered in light of the program's capacity for direct services and patient care. These are by no means mutually exclusive, but it is an important consideration when studying developing interventions (see also Wandersman et al., 2008).

Given these challenges, we suggest that participatory evaluation methods are a useful strategy for assessing programs' work—and doing so in ways that help build programmatic capacity (Cousins & Chouinard, 2012; Torres & Preskill, 2001). In this conceptual replication, we used a participatory evaluation approach to collaborate with local stakeholders to collect data and engage practitioners in the process of evaluating their work (Cousins & Chouinard, 2012). A key advantage of this approach is that program staff become deeply involved in the data collection and analysis process, which tends to increase their interest and investment in the findings (see Cousins & Chouinard, 2012; Cousins & Whitmore, 1998; Patton, 2008; Torres & Preskill, 2001, for reviews). In a traditional research project, the investigators often report back findings (often as a written report) sometime after the work was completed and it can be difficult to mobilize the community to act on the results because they were not involved in the process and/or do not have the interorganizational relationships established to start working on solutions (Fetterman & Wandersman, 2005; Patton, 2008; Wandersman, 2003; Wandersman et al., 2008). By contrast, in participatory evaluation, program staff see the results quickly (because they are often the ones generating the data), and because they have forged connections with other colleagues to make the project happen, it can be easier to launch efforts for change (Cousins & Chouinard, 2012; Cousins & Whitmore, 1998). Indeed, in this project, SANE practitioners were dismayed by the results, and in all six communities, they immediately began working with their colleagues in the legal system on reform strategies (see Campbell et al., 2013, for more details on these postevaluation social change activities). With an intervention model as widespread and rapidly growing as SANE programs, our experience in this project suggests a participatory evaluation paradigm—in conjunction with traditional research approaches—may be an effective strategy for supporting evidence-based practice.

The contributions of this study must be considered in light of three key limitations. First, as noted previously, we did not randomly select programs from the population of all U.S. SANE programs. This was a deliberate decision (not a methodological error) given our concerns about how participation in research/evaluation could have unintended negative consequences on program operations and patient care in new/struggling programs. We specifically sought out SANE programs that had organizational readiness to engage in this kind of work, which is, undoubtedly, a select group within the national population—one that is most likely at a higher level of programmatic functioning. Therefore, the results of this project cannot be generalized to U.S.

SANE programs writ large. Furthermore, we acknowledge that because programs had to apply to be part of this initiative, the sample is self-selected (among programs with organizational readiness), which further limits the generalizability of the findings. Given that our results highlight very high rates of sexual assault case attrition in communities with motivated, engaged, sustainable SANE programs, spanning rural, mid-sized, and urban jurisdictions, there is a critical need for research on what is happening in communities with less interorganizational engagement and commitment to the issue of sexual assault.

Second, we had hoped that all six programs would be able to conduct a pre-SANE/post-SANE design, but in practice only three had the data available for such a design. In addition, program staff were only able to collect 2 (or 3 years) of pre-SANE data, which was then compared with 5-plus years of post-SANE data. The small pre-SANE samples and uneven pre-SANE to post-SANE sample sizes yielded low statistical power, limiting the possibility of detecting significant site-specific findings. Though it would have been ideal to have a comparable number of pre-SANE and post-SANE years, it was impractical, if not impossible because sites simply did not have access to older data, or the older records had been destroyed.

Finally, with respect to the assessment of the dependent variable (criminal justice case outcome), our data collection methods assessed how many cases were "not referred/not charged," which collapses two different scenarios. Cases that were not referred by law enforcement to the prosecutor, and cases that were referred to the prosecutor, but not charged are indistinguishable in these data sets. It certainly is possible to discern whether a case was never referred versus referred/not charged, but doing so requires much more extensive data collection (see Campbell et al., 2009). Such procedures were beyond the scope of what could be asked of and expected from the SANE program sites. In this project, we could determine it was "early" in the criminal justice process, which was sufficient information to guide instrumental change efforts.

In spite of these limitations, this project reveals that SANE programs have improved the criminal justice system's response to sexual assault, but the vast majority of cases are still falling through the cracks very early in the process. These low rates of case referral and charging are occurring in well-established programs that possess organizational readiness and resources, which begs the question, "What is happening in new and/or struggling programs?" Moreover, what is happening in exemplar programs and what are the critical ingredients of their success and how can that information be widely disseminated to practitioners? It is also important to keep in mind that SANE program services are one of only many factors that influence how the criminal justice system processes cases. Deep-seated stereotypes and beliefs about rape are still rampant within the legal system itself (e.g., Page, 2008a, 2008b, 2010) and in the community at large (e.g., Edwards, Turchik, Dardis, Reynolds, & Gidycz, 2011). The findings of this project invite deeper consideration of what SANE programs can do—in and of themselves—to address these issues. As nursing interventions focused on patient care, their mission is first and foremost the well-being of victims (IAFN, 2009); these services may have a positive indirect effect on prosecution rates, but they are

clearly not the sole solution to the problem of underprosecution. The results of this study highlight the importance of a systemic, multidisciplinary approach, as one program or one service is not nearly sufficient in size or scope to tackle the problem of underprosecution.

Authors' Note

Opinions, findings, and conclusions or recommendations expressed in this publication/program/exhibition are those of the author(s) and do not necessarily reflect those of the Department of Justice.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This project was supported by Grant 2009-MU-MU-0002 awarded by the National Institute of Justice, Office of Justice Programs, U.S. Department of Justice.

Notes

1. To clarify the meaning of key terms used in this article, "sexual violence" refers to a broad range of sexually violating events, including "rape" (a nonconsensual act of oral, vaginal, and/or anal penetration committed by the use of force, threat of force, or when an individual is unable to provide consent) and "sexual assault" (a broader range of nonconsensual contact and noncontact sexual offenses, up to and including rape). We use the terms *victim* and *survivor* interchangeably to reflect that sexual assault is a violent crime that takes tremendous strength and courage to survive (see Campbell & Townsend, 2011).
2. Experimental designs have not been used in this literature largely due to ethical concerns about randomly assigning sexual assault victims to a non-SANE condition for medical care. Several studies have indicated that SANE programs are significantly more likely than traditional hospital emergency departments to provide vital services such as sexually transmitted infection (STI) screening and prophylaxis and emergency contraception (Campbell et al., 2006; Ciancone, Wilson, Collette, & Gerson, 2000; Logan, Cole, & Capillo, 2007). As such, the methodological gains of a pure experimental design do not outweigh the potential risks to survivors' health.
3. We were not concerned that many programs did not complete the application. The required elements were not intended to be burdensome by any means, but they did highlight the kind of commitment that would be needed to participate in this project (e.g., organizational support, collaboration with their prosecutors, staff time for some data collection, etc.).
4. In one rural site and one midsized site, their parent hospitals had recently switched to new database systems that unfortunately did not transfer old files in a way that allowed records to be searched by presenting complaint (i.e., sexual assault); as such, they had no way to identify pre-SANE patients. One urban site discovered that pre-SANE records had been lost in a fire.

5. Each site also needed institutional review board (IRB) approval because de-identified data would be later shared with the research team. We had IRB approval from Michigan State University for the multisite project, which provided oversight for our training and technical assistance work with the sites and our analysis of their de-identified data. In other words, each of the six sites had local IRB approval from their home institutions and we had multisite project approval from Michigan State University.
6. The first year of program operation was excluded from the evaluation because the literature on the development of SANE programs has documented that there are often multiple changes in staffing, services, and community relationships during a program's launch (DOJ, 2013). Campbell, Patterson, and Bybee (2012) found that Year 2 cases were most appropriately modeled as the "start" of the program. The most recent year of program services was excluded because it typically takes 1 year (or more) for a case to move through the criminal justice system. Very recent cases would still have pending case outcomes, which could skew the evaluation findings.
7. Campbell et al. (2012) distinguished between cases not referred by police to prosecutors and cases that were referred but not charged by prosecutors. Drawing this distinction requires extensive data collection to cross-check and compare data across police and prosecutor records. In this project, SANE program staff did not have the resources for such intensive efforts. Because the Campbell et al. (2012) study showed that the vast majority of cases that were not prosecuted were due to nonreferral (and therefore the extra work of distinguishing these two categories may not be warranted because one of them was so disproportionately large relative to the other), we decided it would be reasonable to combine categories in this project. In two sites (one urban, one rural), police and prosecutors' records were organized such that it was possible to cross-check this issue fairly easily and, in both sites, program staff confirmed that the vast majority of cases in this combined category were indeed not referred.
8. Overall, the rate of missing data was very low—four of the six sites had no missing data. One urban site had unknown outcomes on 38 of their 382 outcomes (9.9% missing data). One midsized community site had unknown outcomes on 31 of their 232 cases (13.4% missing data). Across all sites, there were 69 cases with unknown outcomes out of a total of 1,765 cases (3.9% missing data). All analyses, both within-site and across-site, were performed on the data with known outcomes.
9. Data from both the post-only design sites (Sites B, D, and F) and the pre-SANE/post-SANE sites (Sites A, C, and E) can be combined in these analyses, provided that one of the variables modeled reflects the differential designs (see below).

References

Black, M. C., Basile, K. C., Breiding, M. J., Smith, S. G., Walters, M. L., Merrick, M. T., et al. (2011). *The National Intimate Partner and Sexual Violence Survey (NISVS): 2010 summary report*. Atlanta, GA: National Center for Injury Prevention and Control, Centers for Disease Control and Prevention.

Campbell, R. (2008). The psychological impact of rape victims' experiences with the legal, medical, and mental health systems. *American Psychologist*, *68*, 702-717.

Campbell, R., Bybee, D., Ford, J. K., Patterson, D., & Ferrell, J. (2009). *A systems change analysis of SANE programs: Identifying the mediating mechanisms of criminal justice system impact* (Final report NIJ Award 2005-WG-BX-0003). Washington, DC: National Institute of Justice.

Campbell, R., & Patterson, D. (2010). *National SANE sustainability project: Final evaluation report*. Enola, PA: National Sexual Violence Resource Center.

Campbell, R., Patterson, D., & Bybee, D. (2012). Prosecution of adult sexual assault cases: A longitudinal analysis of the impact of a Sexual Assault Nurse Examiner program. *Violence Against Women*, *18*, 223-234.

Campbell, R., Patterson, D., & Fehler-Cabral, G. (2010). Using ecological theory to evaluate the effectiveness of an indigenous community intervention: A study of sexual assault nurse examiner (SANE) programs. *American Journal of Community Psychology*, *46*, 263-276.

Campbell, R., Patterson, D., & Lichty, L. F. (2005). The effectiveness of sexual assault nurse examiner (SANE) program: A review of psychological, medical, legal, and community outcomes. *Trauma, Violence, & Abuse*, *6*, 313-329.

Campbell, R., & Townsend, S. M. (2011). Defining the scope of sexual violence. In C. Renzetti, J. Edleson, & R. Bergen (Eds.), *Sourcebook on violence against women* (2nd ed., pp. 95-109). Thousand Oaks, CA: SAGE.

Campbell, R., Townsend, S. M., Bybee, D., Shaw, J. L., & Markowitz, J. (2013). *Implementation of a Sexual Assault Nurse Examiner (SANE) practitioner evaluation toolkit*. Washington, DC: National Institute of Justice.

Campbell, R., Townsend, S. M., Long, S. M., Kinnison, K. E., Pulley, E. M., Adames, S. B., & Wasco, S. M. (2006). Responding to sexual assault victims' medical and emotional needs: A national study of the services provided by SANE programs. *Research in Nursing & Health*, *29*, 384-398.

Ciancone, A., Wilson, C., Collette, R., & Gerson, L. W. (2000). Sexual assault nurse examiner programs in the United States. *Annals of Emergency Medicine*, *35*, 353-357.

Cousins, J. B., & Chouinard, J. A. (2012). *Participatory evaluation up close: An integration of research-based knowledge*. Charlotte, NC: Information Age Publishing.

Cousins, J. B., & Whitmore, E. (1998). Framing participatory evaluation. *New Directions for Evaluation*, *80*, 5-23.

Crandall, C., & Helitzer, D. (2003). *Impact evaluation of a Sexual Assault Nurse Examiner (SANE) program* (Award Number 98-WT-VX-0027). Washington, DC: National Institute of Justice.

Department of Justice. (2013). *A national protocol for sexual assault medical forensic examinations: Adults/adolescents* (2nd ed.). Washington, DC: Author.

Edwards, K. M., Turchik, J. A., Dardis, C. M., Reynolds, N., & Gidycz, C. A. (2011). Rape myths: History, individual and institutional-level presence, and implications for change. *Sex Roles*, *65*, 761-773.

Fetterman, D. M., & Wandersman, A. (2005). *Empowerment evaluation principles in practice*. New York: Guilford Press.

Goldstein, H. (2003). *Multilevel statistical models* (3rd ed.). London, England: Edward Arnold.

Hedeker, D. R., & Gibbons, R. D. (1994). A random-effects ordinal regression model for multilevel analysis. *Biometrics*, *50*, 933-944.

International Association of Forensic Nurses. (2009). *Forensic nurses: Scope and standards of practice*. Silver Springs, MD: American Nursing Association.

International Association of Forensic Nurses. (2013). *Data base of the International Association of Forensic Nurses*. Available from http://www.forensicnurse.org

Kilpatrick, D. G., Resnick, H. S., Ruggiero, K. J., Conoscenti, L. M., & McCauley, J. (2007). *Drug-facilitated, incapacitated, and forcible rape: A national study*. Washington, DC: U.S. Department of Justice.

Logan, T., Cole, J., & Capillo, A. (2007). Sexual Assault Nurse Examiner (SANE) program characteristics, barriers, and lessons learned: A national random sample of programs. *Journal of Forensic Nursing*, *3*, 24-34.

Lonsway, K. A., & Archambault, J. (2012). The "justice gap" for sexual assault cases: Future directions for research and reform. *Violence Against Women*, *18*, 145-168.

Page, A. D. (2008a). Gateway to reform? Policy implications of police officers' attitudes toward rape. *American Journal of Criminal Justice*, *33*, 44-58.

Page, A. D. (2008b). Judging women and defining crime: Police officers' attitudes toward women and rape. *Sociological Spectrum*, *28*, 389-411.

Page, A. D. (2010). True colors: Police officers and rape myth acceptance. *Feminist Criminology*, *5*, 315-334.

Patterson, D., Campbell, R., & Townsend, S. (2006). Sexual Assault Nurse Examiner (SANE) program goals and patient care practices. *Journal of Nursing Scholarship*, *38*, 180-186.

Patton, M. Q. (2008). *Utilization-focused evaluation* (4th ed.). Thousand Oaks, CA: SAGE.

Planty, M., Langton, L., Krebs, C., Berzofsky, M., & Smiley-McDonald, H. (2013). *Female victims of sexual violence, 1994-2010*. Washington, DC: Bureau of Justice Statistics.

Preskill, H., & Boyle, S. (2008). A multidisciplinary model of evaluation capacity building. *American Journal of Evaluation*, *29*, 443-459.

Preskill, H., & Torres, R. T. (2000). *Readiness for organizational learning and evaluation instrument (ROLE)*. Claremont, CA: Author.

Raudenbush, S., Bryk, A., Cheong, Y. F., & Congdon, R. (2011). *HLM 7: Hierarchical linear and nonlinear modeling*. Chicago, IL: Scientific Software International.

Rogers, E. M. (2003). *Diffusion of innovations* (5th ed.). New York: Free Press.

Russ-Eft, D., & Preskill, H. (2001). *Evaluation in organizations: A systematic approach to enhancing learning, performance, and change*. New York: Basic Books.

Schmidt, S. (2009). Shall we really do it again? The powerful concept of replication is neglected in the social sciences. *Review of General Psychology*, *13*, 90-100.

Spohn, C., & Tellis, K. (2012a). The criminal justice system's response to sexual violence. *Violence Against Women*, *18*, 169-192.

Spohn, C., & Tellis, K. (2012b). *Policing and prosecuting sexual assault in Los Angeles city and county: A collaborative study in partnership with the Los Angeles Police Department, the Los Angeles Sheriff's Department, and the Los Angeles County District Attorney's Office* (NCJRS Document 237582). Washington, DC: National Institute of Justice.

Tjaden, P., & Thoennes, N. (2006). *Extent, nature, and consequences of rape victimization: Findings from the National Violence Against Women Survey*. Washington, DC: National Institute of Justice, U.S. Department of Justice.

Torres, R. T., & Preskill, H. (2001). Evaluation and organizational learning: Past, present, and future. *American Journal of Evaluation*, *22*, 387-395.

Wandersman, A. (2003). Community science: Bridging the gap between science and practice within community-centered models. *American Journal of Community Psychology*, *31*, 227-242.

Wandersman, A., Duffy, J., Flaspohler, P., Noonan, R., Lubell, K., Stillman, L., et al. (2008). Bridging the gap between prevention research and practice: The interactive systems framework for dissemination and implementation. *American Journal of Community Psychology*, *41*, 171-181.

Author Biographies

Rebecca Campbell, PhD, is a professor of psychology at Michigan State University. Her research examines how the legal, medical, mental health systems and rape crisis centers respond to the needs of sexual assault victims.

Deborah Bybee, PhD, is a professor of psychology at Michigan State University. She specializes in research design and statistical methodology, focusing on the application of multivariate techniques to understand complex, real-world phenomena, especially those that involve change over time. Much of her research addresses violence against women.

Stephanie M. Townsend, PhD, is president of Townsend Consulting & Evaluation. As a community psychologist, she specializes in developing practice-based evidence in the areas of sexual violence prevention and community responses to sexual violence.

Jessica Shaw, MA, is a doctoral candidate in the Ecological-Community Psychology program at Michigan State University. Her research interests include community responses to sexual assault within the medical and criminal justice systems and program and policy evaluation.

Nidal Karim, PhD, works at CARE as a senior advisor for Gender and Empowerment Impact Measurement. She has also worked as a Behavioral Scientist at the Division for Violence Prevention at CDC

Jenifer Markowitz, ND, RN, WHNP-BC, is a forensic nursing consultant who specializes in issues related to sexual assault and domestic violence, including medical-forensic examinations and professional education and curriculum development. In addition to teaching internationally, Dr. Markowitz provides expert testimony, case consultation, and technical assistance; and develops training materials, resources, and publications for a variety of military and civilian clients.