

# Screening for Intimate Partner Violence in Health Care Settings

## A Randomized Trial

Harriet L. MacMillan, MD, MSc, FRCPC

C. Nadine Wathen, PhD

Ellen Jamieson, MEd

Michael H. Boyle, PhD

Harry S. Shannon, PhD

Marilyn Ford-Gilboe, RN, PhD

Andrew Worster, MD

Barbara Lent, MD

Jeffrey H. Coben, MD

Jacquelyn C. Campbell, PhD

Louise-Anne McNutt, PhD

for the McMaster Violence Against Women Research Group

**F**EW ISSUES IN THE FIELD OF FAMILY violence generate as much controversy as screening women for intimate partner violence (IPV) in health care settings.<sup>1,2</sup> Herein, we use the term *screening* to refer to universal routine inquiry: “a standardized assessment of patients, regardless of their reasons for seeking medical attention,”<sup>1</sup> aimed at identifying women who are experiencing or have recently experienced IPV.

Proponents of screening emphasize the following as a rationale for its implementation: the high prevalence of IPV and associated impairment,<sup>3,4</sup> the high level of acceptability among women about such inquiry,<sup>5,6</sup> the availability of feasible screening techniques,<sup>7,8</sup> and the opportunity to offer support and refer-

**Context** Whether intimate partner violence (IPV) screening reduces violence or improves health outcomes for women is unknown.

**Objective** To determine the effectiveness of IPV screening and communication of positive results to clinicians.

**Design, Setting, and Participants** Randomized controlled trial conducted in 11 emergency departments, 12 family practices, and 3 obstetrics/gynecology clinics in Ontario, Canada, among 6743 English-speaking female patients aged 18 to 64 years who presented between July 2005 and December 2006, could be seen individually, and were well enough to participate.

**Intervention** Women in the screened group (n=3271) self-completed the Woman Abuse Screening Tool (WAST); if a woman screened positive, this information was given to her clinician before the health care visit. Subsequent discussions and/or referrals were at the discretion of the treating clinician. The nonscreened group (n=3472) self-completed the WAST and other measures after their visit.

**Main Outcome Measures** Women disclosing past-year IPV were interviewed at baseline and every 6 months until 18 months regarding IPV reexposure and quality of life (primary outcomes), as well as several health outcomes and potential harms of screening.

**Results** Participant loss to follow-up was high: 43% (148/347) of screened women and 41% (148/360) of nonscreened women. At 18 months (n=411), observed recurrence of IPV among screened vs nonscreened women was 46% vs 53% (modeled odds ratio, 0.82; 95% confidence interval, 0.32-2.12). Screened vs nonscreened women exhibited about a 0.2-SD greater improvement in quality-of-life scores (modeled score difference at 18 months, 3.74; 95% confidence interval, 0.47-7.00). When multiple imputation was used to account for sample loss, differences between groups were reduced and quality-of-life differences were no longer significant. Screened women reported no harms of screening.

**Conclusions** Although sample attrition urges cautious interpretation, the results of this trial do not provide sufficient evidence to support IPV screening in health care settings. Evaluation of services for women after identification of IPV remains a priority.

**Trial Registration** clinicaltrials.gov Identifier: NCT00182468

JAMA. 2009;302(5):493-501

www.jama.com

als to patients once IPV is identified.<sup>6,9</sup>

Organizations such as the US Preventive Services Task Force<sup>10</sup> and the Canadian Task Force on Preventive Health Care<sup>11</sup> have concluded that insufficient evidence exists to recom-

**Author Affiliations** are listed at the end of this article. A complete list of the members of the McMaster Violence Against Women Research Group appears as an eAppendix at <http://www.jama.com>.

**Corresponding Author:** Harriet L. MacMillan, MD, MSc, FRCPC, Offord Centre for Child Studies, Department of Psychiatry and Behavioural Neurosciences, McMaster University, Patterson Bldg, 1200 Main St W, Hamilton, ON L8N 3Z5, Canada (macmilnh@mcmaster.ca).

For editorial comment see p 568.

mend for or against universal screening. The rationale for this position is based on limitations of existing screening approaches including validation studies, insufficient evidence regarding effectiveness of services to which women can be referred once identified, a paucity of evidence that IPV screening improves the likelihood of positive health outcomes, and a lack of studies evaluating the potential harm associated with IPV screening.<sup>10-13</sup>

Nevertheless, clinicians and health care organizations are being encouraged to implement IPV screening. Numerous professional societies recommend routine IPV evaluation, assessment, and/or screening as a part of standard patient care,<sup>8</sup> and the standards of the Joint Commission<sup>14</sup> require that hospitals have objective criteria for identifying and assessing possible victims of abuse and neglect. These policies and recommendations persist despite the lack of evidence regarding effective interventions for IPV. While clinicians are being encouraged to implement routine screening, subsequent actions including referrals are often left to the clinicians' discretion. It is important to understand the potential risks and benefits of screening for IPV in health care settings.

The primary aim of this trial was to examine the effectiveness of IPV screening and communication of a positive screening result to clinicians in health care settings, compared with no screening, in reducing subsequent violence and improving quality of life. We also examined secondary health outcomes for abused women and determined whether screening was associated with harmful consequences.

## METHODS

### Study Settings and Participants

Participants were recruited from 12 primary care sites (family practices and community health centers), 11 acute care sites (emergency departments), and 3 specialty care sites (obstetrics/gynecology clinics) in Ontario, Canada, from July 2005 to December 2006. Individual women were each followed up

for 18 months, starting in July 2005 and ending in July 2008.

### Sample Size and Randomization

Sample size was based on IPV recurrence because this binary measure would require the larger sample size to detect a clinically important difference. Intimate partner violence included emotional, physical, or sexual violence by an intimate partner as measured by the Composite Abuse Scale (CAS). We considered a reduction from a 60% to a 45% recurrence over 18 months to be clinically important. Based on a 2-sided test at the  $\alpha=.05$  level and 80% power, 186 women per group were required.

Randomization was by day or shift (for sites with shifts). A table for each day of the week was created for an 8-week period, and a random number table was used to determine the order of weeks 1 through 8 in the cells. This ensured balance across shifts and days of the week, important because of possible systematic differences in presentation by day or shift. The research coordinator created monthly calendars showing shift allocations for site coordinators.

### Ethical and Safety Considerations

Every effort was made to protect study participants and staff from violence as a consequence of IPV disclosure. Prior to recruitment, clinicians received study-provided, standardized training in responding to IPV. This included an overview of IPV, appropriate responses to IPV disclosure, safety assessment, and information about local community resources for referral. A safety protocol was developed for recruiters and interviewers (available from the authors). A data monitoring committee oversaw the ongoing safety of participants by monitoring primary outcomes and adverse events. The study was approved by the research ethics board of McMaster University/Hamilton Health Sciences and/or the site-specific research ethics board. Written informed consent was obtained from each participant prior to enrollment in the trial.

### Measures

The Woman Abuse Screening Tool (WAST)<sup>15</sup> is an 8-item instrument that measures physical, sexual, and emotional abuse in the last 12 months. It correctly classified 100% of non-abused women and 92% of abused women in a known-group analysis and is highly correlated ( $r=0.96$ ) with the Abuse Risk Inventory.<sup>16</sup> It includes questions such as "Did arguments ever result in hitting, kicking or pushing?" Questions were scored as 0 (never), 1 (sometimes), and 2 (often). Based on data from an earlier trial,<sup>17</sup> a score of 4 or more on the WAST indicates exposure to IPV.

The CAS,<sup>18</sup> a 30-item validated research instrument, was selected as the criterion standard for recurrence of IPV (1 of the 2 primary outcomes) based on its comprehensiveness and psychometric properties. Its 4 subscales correlate highly with corresponding subscales of the Conflict Tactics Scales.<sup>19,20</sup> It includes partner behaviors such as "[My partner] . . . slapped me"; ". . . kept me from medical care"; ". . . harassed me over the telephone"; and ". . . told me I wasn't good enough"; scaled from 0 (never) to 5 (daily). The CAS score is the sum of the 30 items; a score of 7 or higher was the criterion for exposure to IPV. At the initial health care visit, the reference period was the previous 12 months; when administered at 6, 12, and 18 months, it was the previous 6 months.

The second primary outcome, quality of life, was assessed using the World Health Organization Quality of Life (WHOQOL)-Bref instrument, a 26-item measure composed of 4 separate dimensions: physical, psychological, social, and environmental. We decided a priori to use the psychological quality-of-life scale, hypothesizing that it would be most amenable to change. Internal consistency of this scale has ranged from  $\alpha=.86$  to  $\alpha=.91$ .<sup>21,22</sup>

The remaining measures assessed secondary outcomes. Depressive symptoms were assessed using the 20-item Center for Epidemiologic Studies Depression scale,<sup>23</sup> which has a 4-point

scale (0-3) on which the respondent indicates number of times during the past week that a symptom has been experienced. A score of 16 or more was considered a positive screen for depression.

Posttraumatic stress disorder (PTSD) was assessed using the SPAN (Startle, Physiological Arousal, Anger, and Numbness) instrument, a 4-item screen for PTSD validated in part on a sample of abused women.<sup>24</sup> Items were scaled from 0 (not at all distressing) to 4 (extremely distressing). A score of 5 or more was considered a positive screen for PTSD.

Women's alcohol abuse/dependency was assessed with the 5-item TWEAK screening tool. TWEAK is an acronym for "tolerance" (number of drinks to feel high), "worry about drinking," "eye-opener" (morning drinking), "amnesia" (blackouts), and "cut down on drinking."<sup>25</sup> A tolerance of 3 or more drinks scored 2 points, a yes to "worry" scored 2, and the remaining items scored 1 point each. A total score of 2 or more points indicated an alcohol problem. Two questions from the Drug Abuse Severity Test (DAST) assessed previous excess use of prescription drugs and use of street drugs.<sup>26</sup> Endorsing either question indicated a drug problem.

The Short Form 12 health survey, version 2, a valid and reliable short form of the widely used Short Form 36, was used to measure global mental and physical health and well-being. The instrument was designed to measure the effectiveness of interventions and is sensitive to change.<sup>27</sup>

The Consequences of Screening Tool (COST) is a multidimensional questionnaire that measures the effect of being asked IPV screening questions; it was developed for this study.<sup>28</sup> We analyzed the 8-item Effects on Quality of Life subscale because it applied to all women in the screened group, regardless of positive or negative IPV status. This subscale contained the following items, rated on a 5-point scale: (1) "For me, I feel that being asked the questions on partner violence was (good, somewhat good, neither good nor bad, somewhat bad, bad)";

each of the remaining 7 items began, "Because the questions on partner violence were asked . . ." (2) "I feel my home life has become (less difficult . . . more difficult)"; (3) "my feelings about my relationship with my partner are (more positive . . . more negative)"; (4) "I see the quality of my own life as being (better . . . worse)"; (5) "the people in my community who are usually 'there' for me for emotional support are (more available . . . less available)"; (6) "my feelings about myself as a person are (better . . . worse)"; (7) "I feel that the problems in my relationship with my partner are my fault" (disagree . . . agree); and (8) my financial situation has become (better . . . worse)." Items were coded +2 through -2; positive scores indicate benefit while negative scores reflect harm. In a primary care sample of 94 women, the first study to evaluate the COST,<sup>28</sup> this subscale showed a test-retest reliability of 0.74 and an internal consistency of  $\alpha = .73$ ; in this trial, it was  $\alpha = .66$ .

A modified version of the Health and Social Service Utilization questionnaire was used to gather information about women's violence-specific service use, including community-based services.<sup>29</sup>

### Procedures

All women who presented to a study site for a health care visit (the index visit) were approached by a study recruiter to determine eligibility. Patients could be included if they were female, were aged 18 to 64 years, had a male partner at some time in the last 12 months, presented for their own health care visit, were able to separate themselves from those accompanying them, were living within 120 km of the site, were able to speak and read English, were not too ill to participate, and were able to provide informed consent. Eligible women were taken to a private room for the written consent process.

Women were informed that the study was about whether asking all women in health care visits about violence against women does more good than harm. They were told that they might be asked questions about their relationships by completing a form that may be passed

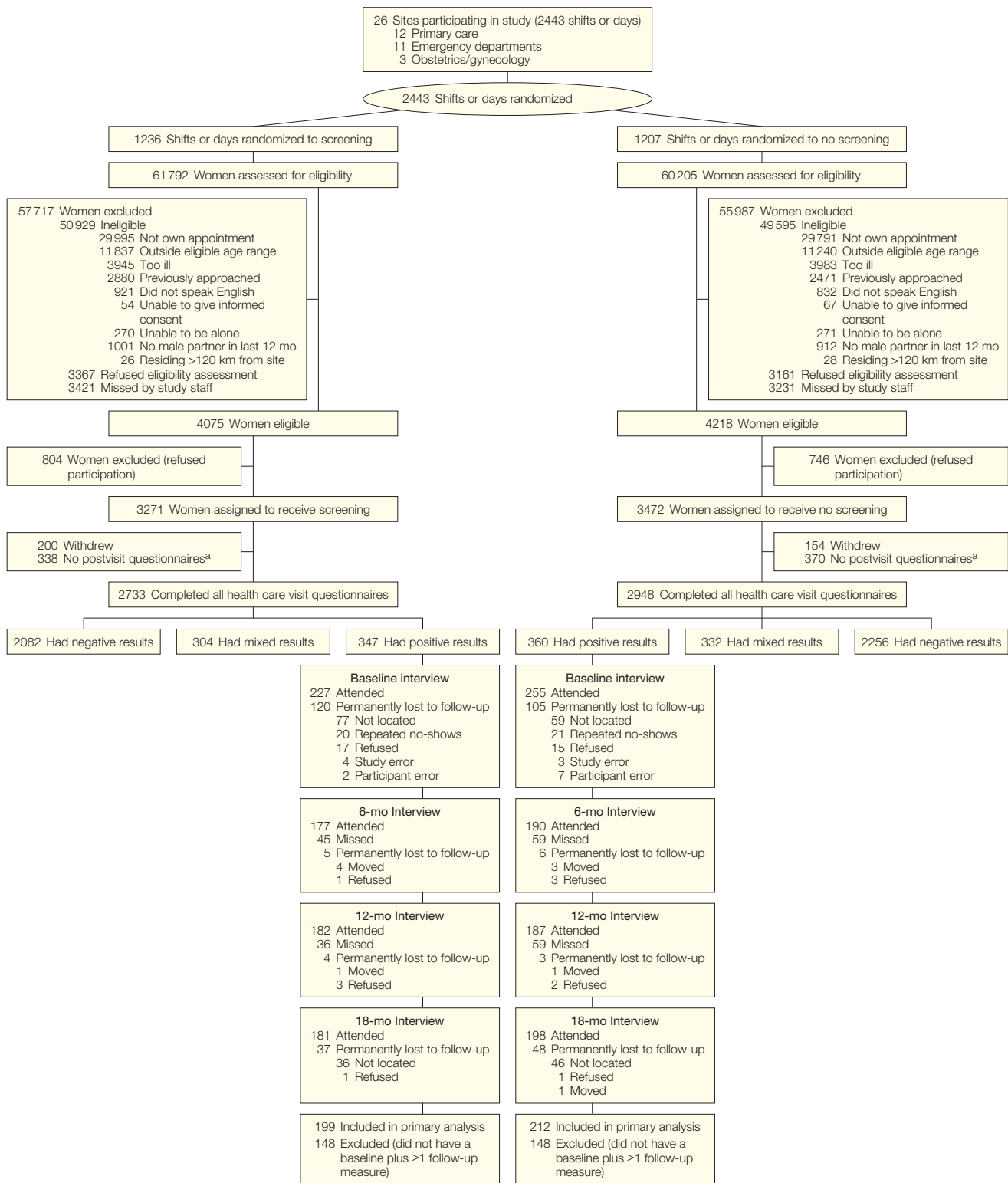
on during this visit to the clinician, who might discuss their situation in more detail. Because witnessing IPV is considered a form of child maltreatment for which the province has mandatory reporting, a statement was included that any information indicating that a child may have experienced harm or was at risk of harm would require a report to child welfare authorities. All women in both groups were then given an information card with contact details for locally available resources for women exposed to violence. This was a business-sized card that listed the names and telephone numbers of local agencies and hotlines for women exposed to violence; it was tailored for each participating community; cards for each community also listed the online service <http://www.shelternet.ca>, which helps women find local shelters and related information and services.

On a screening day, before seeing the clinician, participants answered questions on their socioeconomic circumstances, provided contact information, and completed the WAST screen. For women who screened positive, the research assistant placed the completed WAST questionnaire in the chart for the clinician. Any discussion of positive findings and any further referrals or treatment were left to the discretion of the treating clinician according to his/her usual practice. After their visit, all screened women, regardless of screening results, completed the CAS.

On a no-screening (control) day, before seeing the clinician, participants answered questions on their socioeconomic circumstances and provided contact information. After their clinical encounter, they completed the WAST and the CAS. Clinicians could inquire about abuse if there were clinical indications to do so. In both groups, women who left the site prior to completing the after-visit questionnaires were contacted to complete the questionnaires within a 6-day period.

Interviewers blinded to group assignment met with women within 14 days of the index visit to conduct a baseline interview and again at 6, 12, and 18

**Figure.** Participant Flow



<sup>a</sup>After the health care (index) visit, the screened group completed the Composite Abuse Scale and the nonscreened group completed the Woman Abuse Screening Tool and the Composite Abuse Scale.

months. Participants self-completed the interview package, which included on each occasion all of the primary and secondary outcome measures (with the exception of the CAS and COST, which were completed at the initial health care visit rather than at the baseline interview). Women received graduated honoraria of \$25 to \$50 for baseline through 18-month interviews. To enhance sample retention, brief phone contact was made at 3-month intervals between interviews for the purpose of updating contact information.

### Data Analysis

The statistical software MLwiN<sup>30</sup> was used to model growth trajectories for abuse recurrence (logistic model) and quality of life (linear model), as well as for the secondary measures and for harm. In growth curve analysis, there are 2 steps: first, repeated assessments are modeled as a function of time (number of days since baseline in this study) to estimate individual trajectories that include a starting point (baseline) and change per unit of time or growth (trajectory) for each person; and second, between-group differences (screened vs nonscreened) are estimated at baseline and for rates of change. To account for the study design, with its potential effects of clustering, site and day/shift were included as separate levels in

the analysis and time was allowed to vary between respondents (specified as a random effect). Plotting trajectories of response showed some curvature (ie, a falloff or deceleration of response with time). Accordingly, response was modeled in relation to the square root of time to improve model fit. Results are presented as odds ratios (ORs; screened vs nonscreened) with 95% confidence intervals (CIs) in logistic growth models and score differences (95% CIs) in linear growth models.

Because loss to follow-up was high, average growth measures were estimated from 5 complete files generated through multiple imputation to test the robustness of the observed findings for all enrolled women.

Analyses of all outcomes except harm were run on the randomized sample of 411 women (199 screened and 212 nonscreened) who screened positive on both the WAST and the CAS at the index visit and who had a baseline interview plus at least 1 follow-up interview—the minimum needed to estimate change. The analysis of harm was conducted using baseline data obtained from all women in the screened group who had positive screening results on at least 1 IPV measure and a random selection of those who screened negative on both. This larger sample was used to assess harm because we were interested in the effects

of screening on all women regardless of IPV status; the overwhelming majority of women screened did not report a past-year IPV experience.

### RESULTS

The FIGURE shows sample recruitment and retention. Only 7% (8293/121 997) of women presenting at the health care site were eligible for the study, and 6743 (81%) were randomized to be screened or nonscreened. The 12-month prevalence of IPV at the index visit was 13% (347/2733) in the screened group and 12% (360/2948) in the nonscreened group. In both groups, 11% of women had positive screening results on only 1 instrument; overall, 84% of these were positive on the WAST and negative on the CAS.

Participant loss to follow-up was high: 43% (148/347) in screened women and 41% (148/360) in nonscreened women. A comparison of screened vs nonscreened women by retention status appears in TABLE 1. Women retained in the study had more education, had lower scores on the WAST and CAS, and were less likely to be single. Furthermore, women in the screened group who were lost to follow-up (compared with those retained) reported higher scores on the WAST and CAS.

**Table 1.** Characteristics of Women Retained and Lost to Follow-up by Allocation Status

Characteristics	Retained		Lost		F Ratio	$\chi^2$	P Value
	Screened (n = 199)	Nonscreened (n = 212)	Screened (n = 148)	Nonscreened (n = 148)			
Age, mean (SD), y	33.8 (10.8)	33.9 (10.7)	31.7 (10.1)	32.1 (10.0)	2.03		.11
Education, mean (SD), y	13.7 (2.8)	13.5 (2.8)	13.1 (2.8)	12.9 (2.8)	3.00		.03
WAST score <sup>a</sup>							
Mean (SD)	7.8 (3.0)	8.1 (3.2)	9.2 (3.6)	8.4 (3.5)	8.31		<.001
Median (interquartile range)	19 (11-33)	18 (11-32.8)	28 (13-52.6)	23 (11-42)			
CAS score <sup>a</sup>							
Mean (SD)	26.6 (22.4)	26.5 (22.3)	38.6 (32.7)	31.0 (25.0)	5.21		.001
Median (interquartile range)	8 (5-10)	8 (5.3-10)	9 (6-12)	8 (5-11)			
Single, never married, No. (%)	70 (35)	82 (39)	64 (43)	73 (49)		7.81	.05
Pregnant, No. (%)	8 (4)	18 (9)	12 (8)	15 (10)		5.34	.15
≥1 Child living at home, No. (%)	110 (55)	113 (53)	77 (52)	68 (46)		3.17	.37
Working full- or part-time, No. (%)	108 (54)	111 (52)	73 (49)	74 (50)		1.07	.79
Annual income <Can\$25 000, No. (%)	87 (44)	99 (47)	66 (45)	82 (55)		5.39	.15
Born outside Canada, No. (%)	27 (14)	26 (12)	14 (10)	14 (10)		2.18	.54

Abbreviations: CAS, Composite Abuse Scale; WAST, Woman Abuse Screening Tool.  
<sup>a</sup>See "Methods" section of text for definitions of instrument score ranges.

TABLE 2 shows the observed percentages and mean values for all outcomes and TABLE 3 the results of the growth curve analyses. The trajectory of IPV recurrence risk was downward, with a small, nonsignificant reduction in risk (at 18 months, OR, 0.82; 95% CI, 0.32-2.12) for screened vs nonscreened women. Women in the screened vs nonscreened groups exhibited more rapid improvement in quality of life (at 18 months, 3.74 points higher; 95% CI, 0.47-7.00). Among the secondary measures, only depressive symptoms showed a statistically significant reduction (at 18 months,  $-2.32$ ; 95% CI,  $-4.61$  to  $-0.03$ ).

However, the positive effects associated with quality of life and depressive symptoms are not robust. Estimates derived from multiple imputation were lower for both of these outcomes (2.29 for quality of life and  $-1.97$  for depressive symptoms) and no longer statistically significant.

Our comparison of harm using the COST subscale with the sample that included women exposed and not exposed to IPV (data available from the authors) did not reveal any differences based on exposure status, nor was there any indication of harm associated with screening for either group.

According to self-report immediately after their visit, 44% of screened women and 8% of nonscreened women discussed violence with their clinician during the visit. TABLE 4 shows self-reported use of violence-related services, presented for descriptive purposes only. At each measurement point, contact with family physicians was most frequently reported. The percentages of women in the screened vs nonscreened group who had 1 or more contacts in the 6 months preceding the baseline and 6-, 12-, and 18-month interviews were 75% vs 71%, 67% vs 66%, 63% vs 58%, and 65% vs 64%, respectively.

**Table 2.** Observed Outcomes for Screened vs Nonscreened Women

Outcomes	Baseline		6 mo		12 mo		18 mo	
	Screened (n = 199)	Nonscreened (n = 212)	Screened (n = 177)	Nonscreened (n = 190)	Screened (n = 182)	Nonscreened (n = 187)	Screened (n = 181)	Nonscreened (n = 198)
Intimate partner violence (CAS), No. (%)	199 (100)	212 (100)	104 (59)	122 (64)	91 (50)	110 (59)	83 (46)	105 (53)
Posttraumatic stress disorder screen (SPAN), No. (%)	134 (68)	149 (71)	90 (51)	121 (64)	89 (49)	105 (56)	79 (44)	119 (60)
Alcohol problem (TWEAK), No. (%)	44 (22)	64 (30)	25 (14)	44 (23)	30 (17)	46 (25)	32 (18)	35 (18)
Drug problem (DAST), No. (%)	99 (50)	98 (46)	62 (35)	76 (40)	69 (38)	68 (37)	66 (37)	79 (40)
Quality of life (WHOQOL-Bref), mean (SD) <sup>a</sup>	52.1 (17.8)	50.6 (17.2)	55.1 (17.7)	50.5 (17.9)	56.8 (17.8)	52.5 (18.0)	58.5 (18.4)	52.7 (17.9)
Depression, mean (SD) <sup>a</sup>	25.2 (11.8)	26.4 (11.2)	21.7 (11.2)	25.4 (12.2)	20.7 (11.7)	24.5 (12.6)	21.0 (11.7)	24.4 (12.1)
Physical health (SF-12), mean (SD) <sup>a</sup>	44.4 (11.9)	44.5 (11.7)	46.0 (11.2)	45.4 (11.4)	45.7 (12.2)	45.0 (11.4)	45.7 (12.7)	44.5 (12.0)
Mental health (SF-12), mean (SD) <sup>a</sup>	35.7 (12.1)	34.4 (11.1)	38.8 (11.3)	36.8 (11.3)	39.7 (11.8)	37.1 (12.6)	39.8 (11.4)	36.4 (12.0)

Abbreviations: CAS, Composite Abuse Scale; DAST, Drug Abuse Severity Test; SF-12, Short Form 12; SPAN, Startle, Physiological Arousal, Anger, and Numbness instrument; TWEAK, tolerance/worry about drinking/eye-opener/amnesia/cut down on drinking instrument; WHOQOL-Bref, World Health Organization Quality of Life-Bref instrument.

<sup>a</sup>See "Methods" section of text for definitions of instrument score ranges.

**Table 3.** Model Outcomes Derived From Observed Responses and Multiple Imputation for Screened vs Nonscreened Women

	Odds Ratio (95% CI) <sup>a</sup>				Difference (95% CI) <sup>a</sup>			
	Intimate Partner Violence	PTSD Screening	Alcohol Problem	Drug Problem	Quality of Life	Depression (CES-D)	Physical Health (SF-12)	Mental Health (SF-12)
Observed responses (n = 411)								
6 mo	0.89 (0.51 to 1.54)	0.80 (0.58 to 1.10)	1.14 (0.79 to 1.66)	0.86 (0.63 to 1.17)	2.16 (0.27 to 4.04)	-1.34 (-2.66 to -0.02)	0.88 (-0.37 to 2.13)	0.93 (-0.49 to 2.35)
12 mo	0.85 (0.39 to 1.85)	0.73 (0.47 to 1.15)	1.21 (0.71 to 2.05)	0.80 (0.52 to 1.24)	3.05 (0.39 to 5.72)	-1.90 (-3.77 to -0.03)	1.24 (-0.53 to 3.01)	1.32 (-0.69 to 3.32)
18 mo	0.82 (0.32 to 2.12)	0.68 (0.39 to 1.18)	1.26 (0.66 to 2.40)	0.77 (0.45 to 1.31)	3.74 (0.47 to 7.00)	-2.32 (-4.61 to -0.03)	1.52 (-0.65 to 3.68)	1.61 (-0.84 to 4.07)
Multiple imputation (n = 707)								
6 mo	0.93 (0.61 to 1.41)	0.77 (0.55 to 1.06)	1.13 (0.76 to 1.67)	0.90 (0.59 to 1.36)	1.32 (-0.99 to 3.63)	-1.14 (-2.50 to 0.22)	0.91 (-0.34 to 2.15)	0.60 (-0.98 to 2.19)
12 mo	0.90 (0.50 to 1.63)	0.69 (0.43 to 1.08)	1.18 (0.68 to 2.07)	0.86 (0.48 to 1.55)	1.86 (-1.39 to 5.12)	-1.61 (-3.53 to 0.32)	1.28 (-0.48 to 3.04)	0.85 (-1.39 to 3.09)
18 mo	0.88 (0.43 to 1.82)	0.63 (0.36 to 1.10)	1.23 (0.62 to 2.44)	0.83 (0.41 to 1.71)	2.29 (-1.71 to 6.28)	-1.97 (-4.33 to 0.39)	1.57 (-0.59 to 3.73)	1.05 (-1.70 to 3.79)

Abbreviations: CES-D, Center for Epidemiologic Studies Depression; CI, confidence interval; PTSD, posttraumatic stress disorder; SF-12, Short Form-12.

<sup>a</sup>Adjusted for baseline differences between screened vs nonscreened women.

**Table 4.** Women With At Least 1 Contact With A Specified Service and With Any Service by Group<sup>a</sup>

Service	Baseline		6 mo		12 mo		18 mo	
	Screened (n = 199)	Nonscreened (n = 212)	Screened (n = 177)	Nonscreened (n = 190)	Screened (n = 182)	Nonscreened (n = 187)	Screened (n = 181)	Nonscreened (n = 198)
Family physician	95 (48)	97 (46)	71 (40)	83 (44)	85 (47)	72 (39)	92 (51)	88 (44)
Nurse or nurse practitioner	58 (29)	71 (34)	29 (16)	39 (21)	37 (20)	39 (21)	39 (22)	37 (19)
Emergency department nurse and/or physician	45 (23)	50 (24)	27 (15)	44 (23)	27 (15)	32 (17)	29 (16)	36 (18)
Specialist physician (eg, psychiatrist, gynecologist)	65 (33)	58 (27)	50 (28)	45 (24)	45 (25)	35 (19)	50 (28)	36 (18)
Psychologist or social worker	74 (37)	62 (29)	58 (33)	51 (27)	58 (32)	48 (26)	47 (26)	56 (28)
Crisis phone line	21 (11)	29 (14)	19 (11)	22 (12)	16 (9)	20 (11)	14 (8)	12 (6)
Sexual assault/rape crisis center	7 (4)	3 (1)	5 (3)	5 (3)	4 (2)	5 (3)	5 (3)	5 (3)
Advocacy or counseling service for woman abuse	28 (14)	19 (9)	23 (13)	26 (14)	17 (9)	17 (9)	17 (9)	16 (8)
Women's shelter	20 (10)	22 (10)	11 (6)	13 (7)	9 (5)	10 (5)	7 (4)	9 (5)
Other service	15 (8)	7 (3)	7 (4)	7 (4)	12 (7)	6 (3)	8 (4)	9 (5)
At least 1 service of any type	149 (75)	150 (71)	118 (67)	126 (66)	114 (63)	108 (58)	117 (65)	126 (64)

<sup>a</sup>Data are reported as No. (%).

## COMMENT

In this trial, both the screened and non-screened groups exhibited reductions over time in IPV recurrence, PTSD symptoms, and alcohol problems as well as improvements in quality of life, depression, and mental health.

Women screened using a written, self-completed version of the WAST showed statistically significant improvements in quality-of-life and depression scores compared with those in the non-screened group. However, these differences were relatively small (ie, <0.25 SD) and attenuated to nonsignificance when multiple imputation was used to account for women lost to the analysis. There were no statistically significant differences between groups on other measures, including recurrence of IPV.

In considering explanations for reduction of IPV recurrence and improvements in quality of life in both groups, it is important to note that the study enrolled women during a period of escalated violence that was likely to decrease over time regardless of any intervention (regression toward the mean). Furthermore, self-report is vulnerable to telescoping, leading to an overestimate of IPV in the time frame provided at baseline and susceptible to fatigue bias, producing reduced IPV reporting over time. In addition, since the reference period at the enrollment visit

was the previous 12 months and at later interviews, 6 months, some portion of women may have experienced violence in the early part of the past year, but not in the 6 months before enrollment or thereafter. Finally, it may be that participating in the trial itself—being asked about IPV and its effects—led to benefits.

Why were the differences between groups in IPV recurrence and quality of life not larger? If it is true that study participation conferred benefits, the fact that both groups were interviewed using the same methods at the same intervals would have reduced the likelihood of detecting differences between groups. Screening itself—asking about IPV exposure—may have offered little benefit. As well, every participant received an information card with details about where to seek help in her community. This was necessary to ensure a minimum standard of care for all women participating in the study.

Issues involving violence-related services may have played a role. During training, clinicians were informed of local services available for women exposed to IPV. However, this study was designed to assess effectiveness of screening in the context of existing services and, given the paucity of evidence for IPV services, no specific treatment was implemented for the trial. Second, fewer than

half of the screened-positive women reported having a discussion about violence with their clinician during their visit. Third, self-report data indicate that use of violence-related health and social services was similar between groups; screened women did not access more services. If there were an effective intervention to which women could have been referred once identified as exposed to IPV, it is possible that the results of this trial would have been more positive.

There was no indication that IPV screening was associated with short-term harm among either abused or non-abused women. We and others<sup>10</sup> have raised concerns about possible adverse effects of IPV screening, including reprisal violence, psychological distress, family disruption, and risk of a child being removed from a mother's care following child protective services involvement. We measured harm with an instrument that had not undergone extensive validation prior to its use; however, the primary outcomes confirm that screening did not lead to reprisal violence or decreased quality of life for screened women compared with non-screened women. Consistent with these results, a recent observational study of emergency department patients who screened positive for IPV indicated no safety concerns in the emergency department after undergoing screening.<sup>31</sup>

although only 51% of the sample was interviewed at the 1-week follow-up.

This trial has several methodologic limitations. Among those enrolled in the trial, 42% were lost to follow up, about 30% immediately after the index visit. It is possible that the initial sample loss represents passive refusal by women who may not have been comfortable declining participation in the study when asked by recruiters. Those lost were less educated and less likely to be married, factors related to greater mobility as well as higher risk of violence exposure.<sup>32</sup> Lost participants had higher WAST and CAS scores and, therefore, may have been at higher risk of subsequent violence compared with those retained. This participant loss has important implications for future trials of IPV screening; perhaps screening and follow-up are not feasible, safe, or acceptable to women with high IPV scores. Multiple imputation for lost participants showed that our findings exhibit a bias in favor of the screened group. This supports our conclusion that IPV screening has limited benefits and, perhaps, only for women with moderate IPV scores.

All outcomes were self-reported. We considered the possibility of obtaining health record data but did not consider the quality of such data to be sufficiently high to justify increasing the burden on women in the study by requesting this additional information. Furthermore, although we selected a number of brief self-completed outcome measures to reduce participant burden, the sensitivity of these measures to change may have been more limited than more comprehensive diagnostic measures.

For a number of reasons, this study may not be generalizable and should be replicated. First, the study was conducted in a Canadian province where women have universal access to health care, including follow-up health care services. Second, it was conducted under carefully controlled conditions that may not reflect the reality of most clinical settings. Specifically, highly trained research assistants at every site ensured that women filled out the written screen in

private and that confidentiality was maintained. Research assistants delivered positive screening results to the clinicians immediately following completion of the WAST and maintained efficiency in the screening process. In addition, each clinical site received specific training regarding IPV, including its epidemiology, causes and consequences, and how to have discussions with women regarding this topic. This kind of education regarding IPV is not the norm for many health care settings.<sup>33</sup>

In summary, the results of this study suggest that use of a specific written, self-completed IPV screen with female adult patients presenting to clinical settings within a universal health care system leads to a few modest benefits and is not associated with short-term harms. We used a criterion suggested by others that the threshold for detection change on most quality-of-life measures is about 0.50 SD.<sup>34</sup> Accordingly, we describe our benefits as modest because they exhibit differences less than 0.25 SD. Furthermore, when multiple imputation was used, the differences between groups were no longer statistically significant. We conclude, although sample attrition urges cautious interpretation, that these results do not provide sufficient evidence to support universal IPV screening in health care settings in the absence of an effective intervention to prevent or reduce IPV, especially in the context of the resources required to conduct screening and to deal with the number of women identified by the screening tool.<sup>8,35</sup>

Some might argue that even small improvements in depression or quality-of-life scores justify screening. It is important, however, to consider that of 100 women who underwent screening, 87 women experiencing no abuse (or abuse below the threshold of the criterion standard) were asked about it to identify 13 who disclosed abuse. Second, if the screen alone had been used, as is the case in clinical screening practices, the additional women identified would have led to further effort (and opportunity costs) on the part of the cli-

nicians to clarify whether there was exposure to violence. We carefully considered the threshold at which a woman would be deemed as having a positive screening result; our cut points represented the best balance between sensitivity and specificity found in our earlier study.<sup>17</sup> The specificity of the WAST screen herein was 89%, compared with the earlier observed 96%. In this trial, 535 women (43% of those who screened positive) were identified by the screen, but not the criterion standard, as experiencing IPV. It is critical to balance the number and magnitude of potential benefits of universal screening with the human, opportunity, and resource costs required.

Further research is essential to determine whether these findings are replicated in other settings and samples. It is important to first determine whether screening and follow-up are acceptable or feasible for women reporting exposure to more severe IPV in the past year. We remain concerned about the paucity of evidence for effective treatment and services for women experiencing IPV and recommend that trials evaluating such interventions take high priority. Ideally, future research will determine which interventions are effective in improving health outcomes for abused women, regardless of how they are identified.

**Author Affiliations:** Departments of Psychiatry and Behavioural Neurosciences (Drs MacMillan and Boyle and Ms Jamieson), Pediatrics (Dr MacMillan), Clinical Epidemiology and Biostatistics (Drs Boyle and Shannon), and Emergency Medicine (Dr Worster), McMaster University, Hamilton, Ontario, Canada; Faculty of Information and Media Studies (Dr Wathen), Arthur Labatt Family School of Nursing (Dr Ford-Gilboe), and Department of Family Medicine (Dr Lent), The University of Western Ontario, London, Canada; Department of Emergency Medicine and Community Medicine, West Virginia University School of Medicine, Morgantown (Dr Coben); School of Nursing, Johns Hopkins University, Baltimore, Maryland (Dr Campbell); and Department of Epidemiology, School of Public Health, University at Albany, State University of New York, Albany (Dr McNutt).

**Author Contributions:** Dr MacMillan had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

**Study concept and design:** MacMillan, Wathen, Jamieson, Boyle, Shannon, Ford-Gilboe, Worster, Coben, McNutt.

**Acquisition of data:** MacMillan, Wathen, Worster, Lent, Coben.

**Analysis and interpretation of data:** MacMillan, Wathen, Jamieson, Boyle, Shannon, Ford-Gilboe, Coben, Campbell, McNutt.



*Drafting of the manuscript:* MacMillan, Wathen, Jamieson, Boyle, Ford-Gilboe, Coben.

*Critical revision of the manuscript for important intellectual content:* MacMillan, Wathen, Jamieson, Boyle, Shannon, Ford-Gilboe, Worster, Lent, Coben, Campbell, McNutt.

*Statistical analysis:* MacMillan, Jamieson, Boyle, Shannon, McNutt.

*Obtained funding:* MacMillan, Wathen, Jamieson, Ford-Gilboe, Coben, McNutt.

*Administrative, technical, or material support:* MacMillan, Worster, Lent.

*Study supervision:* MacMillan.

**Financial Disclosures:** None reported.

**Funding/Support:** This research was generously funded by a grant from the former Ontario Women's Health Council (Ontario Ministry of Health and Long-Term Care), whose work was taken on by Echo: Improving Women's Health in Ontario, an agency of the Ontario Ministry of Health and Long-Term Care, in 2007. Dr MacMillan held a Canadian Institutes of Health Research (CIHR) New Emerging Team grant from the Institutes of Gender and Health; Aging; Human Development, Child and Youth Health; Neurosciences, Mental Health, and Addiction; and Population and Public Health and holds the David R. (Dan) Offord Chair in Child Studies. Dr Wathen holds a CIHR-Ontario Women's Health Council New Investigator Award. Dr Boyle holds a Canada Research Chair in the Social Determinants of Child Health.

**Role of the Sponsors:** Neither the Ontario Women's Health Council nor Echo had direct involvement in the design and conduct of the study; in collection, management, analysis, and interpretation of the data; or in preparation, review, or approval of the manuscript.

**Disclaimer:** The views expressed herein do not necessarily reflect those of the Ontario Women's Health Council or Echo. The McMaster Violence Against Women Research Group is solely responsible for the design, implementation, interpretation, and reporting of this study.

**Additional Contributions:** We thank the following members of the data monitoring committee for their time and effort in overseeing the trial: Gene Feder, MD, FRCGP (chair), University of Bristol; Eric Duku, MSc, McMaster University; Kevin Eva, PhD, McMaster University; Michele Hansen, BA, Huron Women's Shelter, Second Stage Housing, and Counselling Services, and Ellen Lipman, MD, MSc, FRCPC, McMaster University. We are very grateful to the participants, clinical staff, and administrators of the participating sites: Brant Community Healthcare System Willett Urgent Care Centre, Brantford Aboriginal Health Centre, Queen West Community Health Centre, Grimsby Medical Associates, Hamilton Aboriginal Health Centre, Hamilton Urban Core Community Health Centre, HART Clinic (Hamilton), Joseph Brant Memorial Hospital Emergency Department, London Family Practice of Bhooma Bhayana, MD, London Health Sciences Centre University Hospital Emergency Department, London Health Sciences Centre Victoria Hospital Emergency Department, McMaster University Medical Centre Emergency Department, Mount Sinai Hospital Emergency Department, Mount Sinai Hospital Obstetrics and Gynecology Department, Niagara Health System Ontario Street Site Prompt Care Centre, North Hamilton Community Health Centre, Rosedale Medical Group (Hamilton), St Joseph's London Family Medical Centre, St Joseph's Health Care London Urgent Care Centre, St Joseph's Health Care London Women's Ambulatory Health Services, St Joseph's Healthcare Hamilton Stoney Creek Campus Urgent Care, Brady Clinic Sudbury, City of Lakes Family Health Team Sudbury, Sudbury Regional Hospital Emergency Department, West End Clinic (Hamilton), and West Lincoln Memorial Hospital Emergency Department. We espe-

cially thank Jill Hancock, RN, Pearl Dodd, Andrew Gulya, BEng, Sandy Brooks, BA, and Susanne King, MSc (all employees of McMaster University at the time of the trial) and our team of interviewers for their efforts in ensuring fidelity throughout all aspects of study procedures. Each of these named contributors was financially compensated from the grant funds for their work with this trial.

**Additional Information:** A complete list of the members of the McMaster Violence Against Women Research Group appears as an eAppendix at <http://www.jama.com>.

## REFERENCES

- Cole TB. Is domestic violence screening helpful? *JAMA*. 2000;284(5):551-553.
- Taket A, Wathen CN, MacMillan HL. Should health professionals screen all women for domestic violence? *PLoS Med*. 2004;1(1):e4.
- American Medical Association. Policy Statement on Family and Intimate Partner Violence: H-515.965. National Advisory Council on Violence and Abuse Policy Compendium. April 2008. [http://www.ama-assn.org/ama1/pub/upload/mm/386/vio\\_policy\\_comp.pdf](http://www.ama-assn.org/ama1/pub/upload/mm/386/vio_policy_comp.pdf). Accessed June 29, 2009.
- Family Violence Prevention Fund. The national consensus guidelines on identifying and responding to domestic violence victimization in health care settings. 2004. [http://endabuse.org/section/programs/health\\_care/\\_consensus\\_guidelines](http://endabuse.org/section/programs/health_care/_consensus_guidelines). Accessed February 27, 2009.
- Rodriguez MA, Bauer HM, McLoughlin E, Grumbach K. Screening and intervention for intimate partner abuse: practices and attitudes of primary care physicians. *JAMA*. 1999;282(5):468-474.
- Phelan MB. Screening for intimate partner violence in medical settings. *Trauma Violence Abuse*. 2007;8(2):199-213.
- Coker AL, Flerx VC, Smith PH, Whitaker DJ, Fadden MK, Williams M. Partner violence screening in rural health care clinics. *Am J Public Health*. 2007;97(7):1319-1325.
- Rabin RF, Jennings JM, Campbell JC, Bair-Merritt MH. Intimate partner violence screening tools: a systematic review. *Am J Prev Med*. 2009;36(5):439-445.e4.
- Chuang CH, Liebschutz JM. Screening for intimate partner violence in the primary care setting: a critical review. *J Clin Outcomes Manag*. 2002;9:565-573.
- Nelson HD, Nygren P, McInerney Y, Klein J; US Preventive Services Task Force. Screening women and elderly adults for family and intimate partner violence: a review of the evidence for the US Preventive Services Task Force. *Ann Intern Med*. 2004;140(5):387-396.
- Wathen CN, MacMillan HL; Canadian Task Force on Preventive Health Care. Prevention of violence against women: recommendation statement from the Canadian Task Force on Preventive Health Care. *CMAJ*. 2003;169(6):582-584.
- Wathen CN, MacMillan HL. Interventions for violence against women: scientific review. *JAMA*. 2003;289(5):589-600.
- Ramsay J, Richardson J, Carter YH, Davidson L, Feder G. Should health professionals screen women for domestic violence? systematic review. *BMJ*. 2002;325(7359):314.
- Joint Commission on Accreditation of Healthcare Organizations. *Comprehensive Manual for Hospitals*. Oakbrook, IL: Joint Commission on Accreditation of Healthcare Organizations; 1996.
- Brown JB, Lent B, Schmidt G, Sas G. Application of the Woman Abuse Screening Tool (WAST) and WAST-Short in the family practice setting. *J Fam Pract*. 2000;49(10):896-903.
- Brown JB, Lent B, Brett PJ, Sas G, Pederson LL. Development of the Woman Abuse Screening Tool for use in family practice. *Fam Med*. 1996;28(6):422-428.
- MacMillan HL, Wathen CN, Jamieson E, et al; McMaster Violence Against Women Research Group. Approaches to screening for intimate partner violence in health care settings: a randomized trial. *JAMA*. 2006;296(5):530-536.
- Hegarty K, Bush R, Sheehan M. The Composite Abuse Scale: further development and assessment of reliability and validity of a multidimensional partner abuse measure in clinical settings. *Violence Vict*. 2005;20(5):529-547.
- Hegarty K, Sheehan M, Schonfeld C. A multidimensional definition of partner abuse: development and preliminary validation of the Composite Abuse Scale. *J Fam Violence*. 1999;14(4):399-415.
- Straus MA. Measuring intrafamily conflict and violence: the Conflict Tactics (CT) Scales. *J Marriage Family*. 1979;41(1):75-88.
- WHOQOL Group. Development of the World Health Organization WHOQOL-BREF quality of life assessment. *Psychol Med*. 1998;28(3):551-558.
- Bonomi AE, Patrick DL, Bushnell DM, Martin M. Validation of the United States' version of the World Health Organization Quality of Life (WHOQOL) instrument. *J Clin Epidemiol*. 2000;53(1):1-12.
- Radloff LS. The CES-D scale: a self-report depression scale for research in the general population. *Appl Psychol Meas*. 1977;1:385-401.
- Meltzer-Brody S, Churchill E, Davidson JR. Derivation of the SPAN, a brief diagnostic screening test for post-traumatic stress disorder. *Psychiatry Res*. 1999;88(1):63-70.
- Russell M, Martier SS, Sokol RJ, et al. Screening for pregnancy risk-drinking. *Alcohol Clin Exp Res*. 1994;18(5):1156-1161.
- El-Bassel N, Schilling RF, Schinke S, et al. Assessing the utility of the Drug Abuse Screening Test in the workplace. *Res Soc Work Pract*. 1997;7:99-114.
- Ware J Jr, Kosinski M, Keller SD. A 12-item short-form health survey: construction of scales and preliminary tests of reliability and validity. *Med Care*. 1996;34(3):220-233.
- Lock J. *The Development of the Consequences of Screening Tool and the Psychometric Assessment of Three Woman Abuse Measures* [thesis]. Hamilton, ON: McMaster University; 2008.
- Browne GB, Arpin K, Corey P, Fitch M, Gafni A. Individual correlates of health service utilization and cost of poor adjustment to chronic illness. *Med Care*. 1990;28(1):43-58.
- Rasbash J, Browne W, Goldstein H, et al. *A Users Guide to MLwiN, Version 2.1*. London, England: Institute of Education; 2000.
- Houry D, Kaslow NJ, Kembal RS, et al. Does screening in the emergency department hurt or help victims of intimate partner violence? *Ann Emerg Med*. 2008;51(4):433-442.
- Wathen CN, Jamieson E, Wilson M, Daly M, Worster A, MacMillan HL. Risk indicators to identify intimate partner violence in the emergency department. *Open Med*. 2007;1(2):e113-e122. <http://www.openmedicine.ca/article/view/63/62>. Accessed February 27, 2009.
- Alpert EJ. Addressing domestic violence: the (long) road ahead. *Ann Intern Med*. 2007;147(9):666-667.
- Norman GR, Sloan JA, Wyrwich KW. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. *Med Care*. 2003;41(5):582-592.
- Wathen CN, Jamieson E, MacMillan HL; McMaster Violence Against Women Research Group. Who is identified by screening for intimate partner violence? *Womens Health Issues*. 2008;18(6):423-432.