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. Cohen, MD
Jacquelyn C. Campbell, PhD
Louise-Anne McNutt, PhD
for the McMaster Violence Against Women Research Group

EW ISSUES IN THE FIELD OF FAMILY violence generate as much controversy as screening women for intimate partner violence (IPV) in health care settings.1,2 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,”3 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,3,4 the high level of acceptability among women about such inquiry,3,6 the availability of feasible screening techniques,7,8 and the opportunity to offer support and referrals to patients once IPV is identified.6,9

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
www.jama.com

©2009 American Medical Association. All rights reserved.
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.10-13

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,8 and the standards of the Joint Commission14 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 α = .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 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)15 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.16 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,17 a score of 4 or more on the WAST indicates exposure to IPV.

The CAS,18 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.19,20 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 α = .86 to α = .91.21,22

The remaining measures assessed secondary outcomes. Depressive symptoms were assessed using the 20-item Center for Epidemiologic Studies Depression scale,23 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. Items were scored 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/dependence 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 drinks to feel high), “worry about acronyms for “tolerance” (number of drinks to feel high), “worry about drinking,” “eye-opener” (morning drinks to feel high), “worry about alcohol”), “amnesia” (blackouts), and “cut down on drinking.” 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. 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.

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. 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, 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.

**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

26 Sites participating in study (2443 shifts or days)
12 Primary care
11 Emergency departments
3 Obstetrics/gynecology

2443 Shifts or days randomized

1236 Shifts or days randomized to screening
61792 Women assessed for eligibility
57717 Women excluded
50929 Ineligible
29995 Not own appointment
11837 Outside eligible age range
3845 Too ill
2880 Previously approached
921 Did not speak English
54 Unable to give informed consent
270 Unable to be alone
1001 No male partner in last 12 mo
26 Residing >120 km from site
3307 Refused eligibility assessment
3421 Missed by study staff

1207 Shifts or days randomized to no screening
60205 Women assessed for eligibility
55967 Women excluded
49595 Ineligible
29701 Not own appointment
11240 Outside eligible age range
3983 Too ill
2471 Previously approached
832 Did not speak English
67 Unable to give informed consent
271 Unable to be alone
912 No male partner in last 12 mo
28 Residing >120 km from site
3161 Refused eligibility assessment
3231 Missed by study staff

4075 Women eligible
804 Women excluded (refused participation)
3271 Women assigned to receive screening
200 Withdrew
338 No postvisit questionnaires

2733 Completed all health care visit questionnaires
2082 Had negative results
304 Had mixed results
347 Had positive results

3472 Women assigned to receive no screening
746 Women excluded (refused participation)
3472 Women assigned to receive no screening
154 Withdrew
370 No postvisit questionnaires

2948 Completed all health care visit questionnaires
360 Had positive results
332 Had mixed results
2256 Had negative results

227 Attended
120 Permanently lost to follow-up
77 Not located
20 Repeated no-shows
17 Refused
4 Study error
2 Participant error

Baseline interview

177 Attended
45 Missed
5 Permanently lost to follow-up
4 Moved
1 Refused

6-mo Interview

182 Attended
36 Missed
4 Permanently lost to follow-up
1 Moved
3 Refused

12-mo Interview

181 Attended
37 Permanently lost to follow-up
36 Not located
1 Refused

18-mo Interview

199 Included in primary analysis
148 Excluded (did not have a baseline plus ≥1 follow-up measure)

212 Included in primary analysis
146 Excluded (did not have a baseline plus ≥1 follow-up measure)

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® was used to model growth trajectories for abuse recurrence (logistic model) and quality of life (linear model), as well as for the secondary outcomes 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

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Retained</th>
<th>Lost</th>
<th>F Ratio</th>
<th>χ²</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, mean (SD), y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Screened (n = 199)</td>
<td>33.8 (10.8)</td>
<td>33.9 (10.7)</td>
<td>31.7 (10.1)</td>
<td>32.1 (10.0)</td>
<td>2.03</td>
</tr>
<tr>
<td>Nonscreened (n = 212)</td>
<td>33.5 (10.8)</td>
<td>33.9 (10.7)</td>
<td>32.1 (10.0)</td>
<td>32.1 (10.0)</td>
<td>3.00</td>
</tr>
<tr>
<td>Education, mean (SD), y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Screened (n = 148)</td>
<td>13.7 (2.6)</td>
<td>13.5 (2.8)</td>
<td>13.1 (2.8)</td>
<td>12.9 (2.8)</td>
<td>2.03</td>
</tr>
<tr>
<td>Nonscreened (n = 148)</td>
<td>13.8 (2.6)</td>
<td>13.5 (2.8)</td>
<td>13.1 (2.8)</td>
<td>12.9 (2.8)</td>
<td>3.00</td>
</tr>
<tr>
<td>WAST score a</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>7.8 (3.0)</td>
<td>8.1 (3.2)</td>
<td>9.2 (3.6)</td>
<td>8.4 (3.5)</td>
<td>8.31</td>
</tr>
<tr>
<td>Median (interquartile range)</td>
<td>19 (11-33)</td>
<td>18 (11-32.8)</td>
<td>28 (13-52.6)</td>
<td>23 (11-42)</td>
<td>5.21</td>
</tr>
<tr>
<td>CAS score a</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>26.6 (22.4)</td>
<td>26.5 (22.3)</td>
<td>38.6 (32.7)</td>
<td>31.0 (25.0)</td>
<td>5.21</td>
</tr>
<tr>
<td>Median (interquartile range)</td>
<td>8 (5-10)</td>
<td>8 (5.3-10)</td>
<td>9 (5-12)</td>
<td>8 (5-11)</td>
<td>7.81</td>
</tr>
<tr>
<td>Single, never married, No. (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Screened (n = 148)</td>
<td>70 (35)</td>
<td>82 (39)</td>
<td>64 (43)</td>
<td>73 (49)</td>
<td>7.81</td>
</tr>
<tr>
<td>Nonscreened (n = 148)</td>
<td>70 (35)</td>
<td>82 (39)</td>
<td>64 (43)</td>
<td>73 (49)</td>
<td>7.81</td>
</tr>
<tr>
<td>Pregnant, No. (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Screened (n = 148)</td>
<td>8 (4)</td>
<td>18 (9)</td>
<td>12 (8)</td>
<td>15 (10)</td>
<td>5.34</td>
</tr>
<tr>
<td>Nonscreened (n = 148)</td>
<td>8 (4)</td>
<td>18 (9)</td>
<td>12 (8)</td>
<td>15 (10)</td>
<td>5.34</td>
</tr>
<tr>
<td>≥1 Child living at home, No. (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Screened (n = 148)</td>
<td>110 (55)</td>
<td>113 (53)</td>
<td>77 (52)</td>
<td>68 (46)</td>
<td>3.17</td>
</tr>
<tr>
<td>Nonscreened (n = 148)</td>
<td>110 (55)</td>
<td>113 (53)</td>
<td>77 (52)</td>
<td>68 (46)</td>
<td>3.17</td>
</tr>
<tr>
<td>Working full- or part-time, No. (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Screened (n = 148)</td>
<td>108 (54)</td>
<td>111 (52)</td>
<td>73 (49)</td>
<td>74 (50)</td>
<td>1.07</td>
</tr>
<tr>
<td>Nonscreened (n = 148)</td>
<td>108 (54)</td>
<td>111 (52)</td>
<td>73 (49)</td>
<td>74 (50)</td>
<td>1.07</td>
</tr>
<tr>
<td>Annual income &lt; Can $25 000, No. (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Screened (n = 148)</td>
<td>87 (44)</td>
<td>99 (47)</td>
<td>66 (45)</td>
<td>82 (55)</td>
<td>5.39</td>
</tr>
<tr>
<td>Nonscreened (n = 148)</td>
<td>87 (44)</td>
<td>99 (47)</td>
<td>66 (45)</td>
<td>82 (55)</td>
<td>5.39</td>
</tr>
<tr>
<td>Born outside Canada, No. (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Screened (n = 148)</td>
<td>27 (14)</td>
<td>26 (12)</td>
<td>14 (10)</td>
<td>14 (10)</td>
<td>2.18</td>
</tr>
</tbody>
</table>

Abbreviations: CAS, Composite Abuse Scale; WAST, Woman Abuse Screening Tool.

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, 2.12) for screened vs nonscreened women, 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 73% vs 71%, 67% vs 66%, 63% vs 58%, and 65% vs 64%, respectively.

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

<table>
<thead>
<tr>
<th>Outcomes</th>
<th>Baseline (n = 212)</th>
<th>6 mo (n = 177)</th>
<th>12 mo (n = 190)</th>
<th>18 mo (n = 187)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intimate partner violence (CAS), No. (%)</td>
<td>199 (100)</td>
<td>212 (100)</td>
<td>104 (59)</td>
<td>122 (64)</td>
</tr>
<tr>
<td>Posttraumatic stress disorder screen (SPAN), No. (%)</td>
<td>134 (68)</td>
<td>149 (71)</td>
<td>90 (51)</td>
<td>121 (64)</td>
</tr>
<tr>
<td>Alcohol problem (TWEAK), No. (%)</td>
<td>44 (22)</td>
<td>64 (30)</td>
<td>25 (14)</td>
<td>44 (23)</td>
</tr>
<tr>
<td>Drug problem (DAST), No. (%)</td>
<td>99 (50)</td>
<td>98 (46)</td>
<td>62 (35)</td>
<td>76 (40)</td>
</tr>
<tr>
<td>Quality of life (WHOQOL-Bref), mean (SD)</td>
<td>52.1 (17.8)</td>
<td>50.6 (17.2)</td>
<td>55.1 (17.7)</td>
<td>50.5 (17.9)</td>
</tr>
<tr>
<td>Depression, mean (SD)</td>
<td>25.2 (11.8)</td>
<td>26.4 (11.2)</td>
<td>29.4 (12.2)</td>
<td>27.0 (12.6)</td>
</tr>
<tr>
<td>Physical health (SF-12), mean (SD)</td>
<td>44.4 (11.9)</td>
<td>44.5 (11.7)</td>
<td>46.0 (11.2)</td>
<td>45.4 (11.4)</td>
</tr>
<tr>
<td>Mental health (SF-12), mean (SD)</td>
<td>35.7 (12.1)</td>
<td>34.4 (11.1)</td>
<td>38.8 (11.3)</td>
<td>36.8 (11.3)</td>
</tr>
</tbody>
</table>

Abbreviations: CAS, Composite Abuse Scale; DAST, Drug Abuse Severity Test; SF-12, Short Form–12; SPAN, Startle, Physiological Arousal, Anger, and Numbness instrument; WHOQOL-Bref, World Health Organization Quality of Life–Bref instrument.

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

<table>
<thead>
<tr>
<th>Odds Ratio (95% CI)</th>
<th>Difference (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intimate Partner Violence</td>
<td>PTSD Screening</td>
</tr>
<tr>
<td>Observed responses (n = 411)</td>
<td>0.89 (0.51 to 1.54)</td>
</tr>
<tr>
<td>6 mo</td>
<td>0.85 (0.39 to 1.85)</td>
</tr>
<tr>
<td>12 mo</td>
<td>0.82 (0.32 to 2.12)</td>
</tr>
<tr>
<td>18 mo</td>
<td>0.93 (0.61 to 1.41)</td>
</tr>
<tr>
<td>Multiple imputation (n = 707)</td>
<td>0.90 (0.50 to 1.63)</td>
</tr>
<tr>
<td>6 mo</td>
<td>0.88 (0.43 to 1.82)</td>
</tr>
<tr>
<td>12 mo</td>
<td>0.90 (0.43 to 1.82)</td>
</tr>
<tr>
<td>18 mo</td>
<td>0.90 (0.43 to 1.82)</td>
</tr>
</tbody>
</table>

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

*Adjusted for baseline differences between screened vs nonscreened women.
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 abused women. We and others \(^1\) 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 was associated with short-term harm among either abused or non-abused women. We and others \(^1\) 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 was associated with short-term harm among either abused or non-abused women. We and others \(^1\) 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 was associated with short-term harm among either abused or non-abused women.

<table>
<thead>
<tr>
<th>Service</th>
<th>Baseline</th>
<th>6 mo</th>
<th>12 mo</th>
<th>18 mo</th>
</tr>
</thead>
<tbody>
<tr>
<td>Family physician</td>
<td>Screened (n = 199)</td>
<td>95 (48)</td>
<td>97 (46)</td>
<td>71 (40)</td>
</tr>
<tr>
<td>Nurse or nurse practitioner</td>
<td>Screened (n = 177)</td>
<td>58 (29)</td>
<td>71 (34)</td>
<td>29 (16)</td>
</tr>
<tr>
<td>Emergency department nurse and/or physician</td>
<td>Screened (n = 182)</td>
<td>45 (23)</td>
<td>50 (24)</td>
<td>27 (15)</td>
</tr>
<tr>
<td>Specialist physician (eg, psychiatrist, gynecologist)</td>
<td>Screened (n = 181)</td>
<td>65 (33)</td>
<td>58 (27)</td>
<td>50 (28)</td>
</tr>
<tr>
<td>Psychologist or social worker</td>
<td>Screened (n = 180)</td>
<td>74 (37)</td>
<td>62 (29)</td>
<td>58 (33)</td>
</tr>
<tr>
<td>Crisis phone line</td>
<td>Screened (n = 179)</td>
<td>21 (11)</td>
<td>29 (14)</td>
<td>19 (11)</td>
</tr>
<tr>
<td>Sexual assault/rape crisis center</td>
<td>Screened (n = 178)</td>
<td>7 (4)</td>
<td>3 (1)</td>
<td>5 (3)</td>
</tr>
<tr>
<td>Advocacy or counseling service for woman abuse</td>
<td>Screened (n = 177)</td>
<td>28 (14)</td>
<td>19 (9)</td>
<td>23 (13)</td>
</tr>
<tr>
<td>Women’s shelter</td>
<td>Screened (n = 176)</td>
<td>20 (10)</td>
<td>22 (10)</td>
<td>11 (6)</td>
</tr>
<tr>
<td>Other service</td>
<td>Screened (n = 175)</td>
<td>15 (8)</td>
<td>7 (3)</td>
<td>7 (4)</td>
</tr>
<tr>
<td>At least 1 service of any type</td>
<td>Screened (n = 174)</td>
<td>149 (75)</td>
<td>150 (71)</td>
<td>118 (67)</td>
</tr>
</tbody>
</table>

\(^{1}\)Data are reported as No. (%).
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.21 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.33

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.34 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.35

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 clinicians 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.37 The specificity of the WAST screen herein was 89%, compared with the earlier observed 90%. 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 Behavioral Neurosciences (Dr 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.
We thank the following Additional Contributions: The views expressed herein do not necessarily reflect the opinions of the McMaster Violence Against Women Research Group. Approaches to screening for intimate partner violence in health care settings: a randomized trial. JAMA. 2006;295(6):530-536.


©2009 American Medical Association. All rights reserved.