The effectiveness of interventions during and after residence in women’s shelters: a meta-analysis

Irene E. Jonker¹, Marit Sijbrandij²,³, Maurice J. A. van Luijtelaar¹, Pim Cuijpers³, Judith R. L. M. Wolf¹

¹ Department of Primary and Community Care, Radboud university medical center, The Netherlands
² Department of Clinical Psychology, VU University Amsterdam, The Netherlands
³ EMGO Institute for Health and Care Research, The Netherlands

Correspondence: Irene Jonker, Department of Primary and Community Care, Radboud university medical center, Impuls, Netherlands Center for Social Care Research. P.O. Box 9101, 6500 HB Nijmegen, The Netherlands, Tel: + 31 24 361 44 20, Fax: + 31 24 361 95 61, e-mail: irene.jonker@radboudumc.nl

Introduction

Every year, ~300,000 women and children in the United States flee to women’s shelters due to intimate partner violence (IPV).¹ European estimates, for instance, from the United Kingdom and the Netherlands indicate percentages that are comparable.² The primary purpose of shelters is to provide a safe haven to stop the abuse and prevent further harm.³ Other goals are to improve (mental) health outcomes, such as improving quality of life and reducing symptoms of PTSD and depression. Social outcomes of shelter interventions include increasing access to resources⁴ and increasing social support.

A well-known intervention administered in shelters is advocacy counselling (see Sullivan⁵–⁷), which focuses on women’s safety, their quality of life and their access to community resources. Female advocates assist women for a period of 10 weeks, in gaining access to resources (e.g., obtaining material goods, education, finances, legal assistance, healthcare, social support, employment and issues enclosing their children).

A social support intervention may also be administered, such as providing information on resources to women and an environment to chat with a counsellor or friends.⁸

A different type of intervention, which is not given to all women, but to a selected sample of shelter-based abused women with symptoms of psychological distress, is cognitive–behavioural therapy (CBT). Most CBT interventions offered in shelters are focused on reducing symptoms of psychological distress as a result of potential traumatic experiences, such as PTSD symptoms.⁹ During CBT, maladaptive behaviours and cognitive processes are addressed through protocoded systematic procedures.⁹

It is unclear to what extent shelter services are efficacious in reducing re-abuse and improving health outcomes in shelter-based abused women. Previous systematic reviews were not specifically aimed to analyse interventions offered by shelters, but included interventions targeted to the broader category of IPV victims.¹⁰¹¹ In the present study, we conducted a meta-analysis to study the effectiveness of all types of psychological and psychosocial interventions administered to female IPV victims during or after their stays in shelters on mental health, abuse, and social outcomes.

Methods

Study search


Study selection

Selected studies met the following inclusion criteria. Participants were female IPV victims (aged ≥18 years), recruited through shelters. The intervention was carried out during and/or after shelter residence. As a scan of the literature suggested that the number of randomized controlled trials (RCTs) would be insufficient for the analysis, we also included non-randomized controlled studies. Control groups received care-as-usual (services typically provided at the same setting or with minimal additions). The following outcomes were considered: (i) mental health (symptoms...
of depression, PTSD or anxiety, quality of life); (ii) abuse (physical and/or sexual abuse); or (iii) social (social support; ability to obtain resources, health care utilization).

Quality assessment and data extraction

Titles and abstracts were reviewed independently by two researchers according to the study selection criteria. Disagreement was solved after discussion. Two independent researchers independently assessed the quality of each study and classified them as good, fair, or poor, based on the Criteria for Grading Internal Validity developed by the US Preventive Services Task Force (USPSTF). Because less rigorous (quasi-experimental) study designs were allowed in the current review, a modification of this rating system was used. This included clear definitions of intervention; use of a control group; randomized allocation to groups; initial resemblance between groups; a retention rate >80%; maintenance of comparable groups; a sample size ≥50; all important outcomes considered; equal, valid and reliable measurements; follow-up measurements and intention-to-treat analysis for evaluating the internal validity of different study designs.

Data were extracted on mental health outcomes (e.g., clinical-assessed or self-reported depression or PTSD severity, quality of life), abuse outcomes (e.g., incidence and/or severity of re-abuse) and social outcomes (e.g., perceived social support and the effectiveness in obtaining resources).

Analysis

We calculated mean effect sizes for each outcome category (mental health, re-abuse and social outcomes) using Comprehensive Meta-Analysis, version 2.2.021. We used intention-to-treat sample sizes. In studies that included more than one post-intervention assessment, data from the first measurement were taken.

Effect sizes were calculated by subtracting the post-treatment mean score of the intervention group from the mean score of the control group, and dividing the result by the pooled standard deviations of the two groups. Effect sizes corresponding to the dichotomous outcome variable of re-abuse incidence were transformed into Hedges’ g according to the procedures implemented in Comprehensive Meta-Analysis. All effect sizes were adjusted for small sample bias according to the procedures suggested by Hedges and Olkin (Hedges’ g). We calculated an average Hedges’ g effect size for studies that included multiple outcome measures within each outcome category (mental health, abuse, social outcomes). Effect sizes of ≥0.8 were considered large, effect sizes of 0.79–0.5 moderate and effect sizes of 0.49–0.2 small.

Because we expected considerable heterogeneity, we calculated mean effect sizes using the random effects model. As an indicator of heterogeneity, we calculated the I²-statistic. An I²-value of 0% indicates no observed heterogeneity, and larger values show increasing heterogeneity, with 25% regarded as low, 50% as moderate and 75% as high heterogeneity. We calculated 95% confidence intervals around I² using the non-central chi-squared-based approach within the heterogi module for Stata. We also calculated the Q-statistic, but only report whether this was significant.

Hedges’ g is not easy to interpret from a clinical perspective. Therefore, we transformed the standardized mean differences into the numbers-needed-to-treat (NNT). The NNT indicates the number of patients that have to be treated in order to generate one additional positive outcome.

Subgroup analyses were conducted according to the mixed-effect model. In this model, studies within subgroups are pooled with the random effects model, while tests for significant differences between subgroups are conducted with the fixed-effects model. For continuous variables, we used meta-regression analyses to test whether there was a significant relationship between the continuous variable and the effect size, as indicated with a Z-value and an associated P-value.

Publication bias was tested by inspecting the funnel plot on primary outcome measures and by Duval and Tweedie’s trim and fill procedure, which yields an estimate of the effect size after the publication bias has been taken into account (as implemented in Comprehensive Meta-analysis, version 2.2.021). We also conducted Egger’s test of the intercept to quantify the bias captured by the funnel plot and to test whether it was significant.

Power calculation

Because we expected only a limited number of studies, we conducted a power calculation, according to the procedures described by Borenstein and colleagues, to examine how many studies would have to be included in order to have sufficient statistical power to identify relevant effects. We hoped to find a sufficient number of studies to be able to identify a moderate effect size of 0.5. These calculations indicated that we would need to include at least six studies with a mean sample size of 50 (25 participants per condition), to be able to detect an effect size of d = 0.50 (conservatively assuming a high level of between-study variance, r², a statistical power of 0.80 and a significance level, alpha, of 0.05). Alternatively, we would need four studies with 80 participants each to detect an effect size of d = 0.50, or three studies with 100 participants.

Results

Selected studies

The search yielded 2989 citations (see supplemental material for the search strategy of one of the databases). A total of 86 articles matched the inclusion criteria and were retrieved in full length for further review. Fifteen of these satisfied appraisal criteria, including hand-searched relevant references cited in key articles and references from Google or Google Scholar. The 15 articles described 10 unique intervention studies. A list of excluded studies can be obtained from the authors. The supplementary figure S1 displays the results of the systematic search.

Study characteristics

Ten studies, published between 1991 and 2011, were included in the meta-analysis, involving 726 participants (375 intervention participants and 351 controls; supplementary table S1). Eight studies were RCTs, two were non-randomized controlled trials. Nine studies were carried out in the United States and one in South Korea. Characteristics of study samples varied across the studies. Most studies included unselected samples of residents or former residents of shelters who had reported IPV. One study only included women with subthreshold or full PTSD, and two studies were confined to women who had a child with a conduct disorder or oppositional-defiant disorder. In five studies, the intervention was carried out during shelter residence, and the other five examined post-shelter interventions. Different intervention types were studied, including a social support intervention, group counselling aimed at improving assertiveness, communication and problem-solving, a music therapy, an intervention for subthreshold or full PTSD symptoms including cognitive behavioural techniques, an intervention for mothers and their children with conduct disorder focused on increasing safety and on teaching parenting skills, group counselling to improve stress management and coping skills and brief one-on-one advocacy services. Seven interventions were individual interventions, and three were group interventions.

Publication bias was tested by inspecting the funnel plot on primary outcome measures and by Duval and Tweedie’s trim and fill procedure, which yields an estimate of the effect size after the publication bias has been taken into account (as implemented in Comprehensive Meta-analysis, version 2.2.021). We also conducted Egger’s test of the intercept to quantify the bias captured by the funnel plot and to test whether it was significant.
The outcomes assessed varied across studies, although most studies assessed mental health outcomes, most notably measures of depression or of general psychological distress. Abuse outcomes (incidence of physical re-abuse) were measured in five studies. Three studies measured social outcomes, e.g., social support or the obtaining of resources.

Quality of included studies and risk of bias

The quality of the 10 included studies varied. Judged by the pre-established criteria for the current review, six studies received a quality rating of 'fair', four studies had a 'poor' rating, and no 'good' rating was assigned. Poor or fair ratings may be attributed in particular to small sample sizes (< 50), absence of randomization, lack of follow-up, no intention-to-treat analysis, low retention rate (<80%), and lack of information about blinding of interviewers.

Mental health outcomes

Nine studies, including 685 participants (350 intervention participants, 335 controls), examined mental health outcomes. We could compare the effects of interventions provided during and after shelter stay with control group data. We found a significant mean overall effect for mental health outcomes of 0.39 (95% CI: 0.24–0.54; Z = 5.13, P < .001), which corresponds with a NNT of 4.59. Heterogeneity was low (I² = 0.0%, ns.). The effect sizes and 95% confidence intervals of each study are presented in figure 1. A post-hoc power calculation showed that the statistical power of these nine studies was 0.999. Results of the meta-analysis are reported in table 1.

Abuse outcomes

Five studies, including 553 participants (284 intervention participants and 269 controls) measured abuse outcomes. The effect size was 0.32 (95% CI: 0.08–0.55; NNT= 5.95), which was significant (P < 0.01). Heterogeneity was low (22.8%) although the 95% confidence interval was broad (0–68%). The effect sizes and 95% confidence intervals of each study are presented in figure 2. A post-hoc power calculation showed that the statistical power of these five studies was 0.73.

Social outcomes

Five studies, with 541 participants (279 intervention participants and 262 controls), included social outcomes (see figure 3 for effect sizes and 95% confidence intervals). The effect size was 0.71 (95% CI: 0.54–0.88; NNT= 2.60), which was significant (P < 0.001). Heterogeneity between the studies was low (I² = 0.0%), and the 95% confidence interval was broad (0–79). A post-hoc power calculation showed that the statistical power of these five studies was 0.999.

The effectiveness of women’s shelter interventions

We found some indications for publication bias. After adjustment for publication bias according to Duval and Tweedie’s trim and fill procedure, the effect size dropped from 0.39 to 0.29 (95% CI: 0.12–0.47; number of trimmed studies: 4). Egger’s test was not significant. Subgroup analyses for mental health outcomes did not indicate that the effect size differed significantly in randomized trials compared with non-randomized trials (P > 0.1), but there was a trend indicating that studies in which the intervention was delivered during shelter had a higher effect size than studies in which the intervention was given after the shelter (P < 0.1).
The aim of the present study was to analyse the effectiveness of interventions for female victims of IPV as administered during and/or after their residence in a shelter in terms of mental health, re-abuse and social outcomes. This review identified 10 studies (reported in 15 papers), in which 726 participants had been recruited via shelters. Nine of the 10 included studies assessed mental health outcomes, specifically depression symptoms, PTSD symptoms, anxiety and/or general psychological distress. Only five assessed abuse outcomes, most notably physical re-abuse incidence. A similar number of studies assessed social outcomes such as social support and obtaining resources. We found that interventions, whether provided during or after a stay in a shelter, significantly improved mental health, abuse and social outcomes as compared to control interventions (largely care as usual). Our findings extend the previous meta-analysis on the effectiveness of advocacy interventions of Ramsey and colleagues. Subgroup analyses showed that intensive advocacy aimed at individual women may help to terminate abuse, and improve quality of life up to two years after the intervention. However, it could not be determined yet whether they had a beneficial effect on mental health outcomes, such as depression, quality of life and psychological distress. The current study found evidence for benefits of shelter-based interventions on these outcomes. Moreover, the Ramsey et al. meta-analysis was limited to advocacy interventions aimed at individual women and did not include other types of interventions such as family-based interventions, which may also be provided in women’s shelters. Finally, we included two recent studies that were not included in the Ramsay et al. meta-analysis.

It is important to acknowledge that there is currently no consensus on what outcomes are primary in interventions for shelter-based abused women, and not even about whether re-abuse incidence is an appropriate measure for evaluating IPV interventions. This is reflected in the wide variety of outcome measures included in the studies in this meta-analysis. Because it has previously been found that shelter-based abused women report needs in several domains, including mental health, safety and psychological needs, we grouped the outcomes in this meta-analysis accordingly. We propose that future RCTs evaluating shelter interventions include fewer, but well-validated, standardized measures for each of these three domains (see also Ramsay et al.). With respect to the social domain, we suggest that in addition to perceived social support, it may also be important to include measures for access to resources (such as finding housing or a job). Indeed, being financial autonomous protects women and enables them to leave abusive relationships, at least in the more developed countries.

The results of our meta-analysis should be interpreted with caution. As with all meta-analyses, the strength of the overall evidence is limited by the number and quality of the studies included. The 10 identified studies were few in number and very diverse in terms of type, goals, content, duration and intensity of the interventions. Also, the quality of the included studies was suboptimal: four studies were rated as poor and six as fair. Quality ratings were based on the information provided in the articles, but relevant information about the content of the interventions, fidelity assessments, the reporting of missing data and loss to follow-up was often lacking. In addition, due to the small number of studies, subgroups analyses could only be conducted for mental health outcomes, but not for abuse and social outcomes. Finally,

### Table 1 Effects of interventions during and after residence in women's shelters: Hedges' $g^a$

<table>
<thead>
<tr>
<th>Outcomes</th>
<th>N</th>
<th>$g$</th>
<th>95% CI</th>
<th>$I^2$</th>
<th>95% CI$^b$</th>
<th>NNT</th>
<th>$P^c$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mental health</td>
<td>9</td>
<td>0.39</td>
<td>0.24-0.54***</td>
<td>0.0</td>
<td>0-65</td>
<td>4.59</td>
<td></td>
</tr>
<tr>
<td>Abuse</td>
<td>5</td>
<td>0.32</td>
<td>0.08-0.55**</td>
<td>22.8</td>
<td>0-68</td>
<td>5.95</td>
<td></td>
</tr>
<tr>
<td>Social</td>
<td>5</td>
<td>0.71</td>
<td>0.54-0.88***</td>
<td>0.0</td>
<td>0-79</td>
<td>2.60</td>
<td></td>
</tr>
<tr>
<td>Subgroup analyses (for mental health)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time of help</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>During shelter</td>
<td>5</td>
<td>0.61</td>
<td>0.28-0.94***</td>
<td>0</td>
<td>0-82</td>
<td>2.99</td>
<td>0.12</td>
</tr>
<tr>
<td>Post-shelter</td>
<td>4</td>
<td>0.32</td>
<td>0.14-0.49***</td>
<td>12.3</td>
<td>0-85</td>
<td>5.56</td>
<td></td>
</tr>
<tr>
<td>Design</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No RCT</td>
<td>2</td>
<td>0.19</td>
<td>-0.35-0.73%</td>
<td>0</td>
<td>d</td>
<td>9.43</td>
<td>0.42</td>
</tr>
<tr>
<td>RCT</td>
<td>7</td>
<td>0.42</td>
<td>0.25-0.60***</td>
<td>11.6</td>
<td>0-74</td>
<td>4.27</td>
<td></td>
</tr>
</tbody>
</table>

*P<0.05; **P<0.01; ***P<0.001.

<table>
<thead>
<tr>
<th>Study name</th>
<th>Outcome</th>
<th>Statistics for each study</th>
<th>Hedges's g and 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Hedges's g</td>
<td>Lower limit</td>
</tr>
<tr>
<td>Constantino, Kim &amp; Crane (2005)</td>
<td>S</td>
<td>0.84</td>
<td>0.03</td>
</tr>
<tr>
<td>Johnson, Zlotnick &amp; Perez (2011)</td>
<td>Combined</td>
<td>0.44</td>
<td>-0.03</td>
</tr>
<tr>
<td>Sullivan &amp; Bybee (1999)</td>
<td>Combined</td>
<td>0.68</td>
<td>0.43</td>
</tr>
<tr>
<td>Sullivan (1991)</td>
<td>S</td>
<td>0.81</td>
<td>0.17</td>
</tr>
<tr>
<td>Sullivan et al (1992)</td>
<td>S</td>
<td>0.86</td>
<td>0.51</td>
</tr>
</tbody>
</table>

Figure 3 Effect size measures (Hedges’ $g$) comparing shelter interventions to control for social (S) outcomes.

### Discussion

The aim of the present study was to analyse the effectiveness of interventions for female victims of IPV as administered during and/or after their residence in a shelter in terms of mental health, re-abuse and social outcomes. This review identified 10 studies (reported in 15 papers), in which 726 participants had been recruited via shelters. Nine of the 10 included studies assessed mental health outcomes, specifically depression symptoms, PTSD symptoms, anxiety and/or general psychological distress. Only five assessed abuse outcomes, most notably physical re-abuse incidence. A similar number of studies assessed social outcomes such as social support and obtaining resources. We found that interventions, whether provided during or after a stay in a shelter, significantly improved mental health, abuse and social outcomes as compared to control interventions (largely care as usual). Our findings extend the previous meta-analysis on the effectiveness of advocacy interventions of Ramsey and colleagues. Subgroup analyses showed that intensive advocacy aimed at individual women may help to terminate abuse, and improve quality of life up to two years after the intervention. However, it could not be determined yet whether they had a beneficial effect on health outcomes, such as depression, quality of life and psychological distress. The current study found evidence for benefits of shelter-based interventions on these outcomes. Moreover, the Ramsey et al. meta-analysis was limited to advocacy interventions aimed at individual women and did not include other types of interventions such as family-based interventions, which may also be provided in women’s shelters. Finally, we included two recent studies that were not included in the Ramsay et al. meta-analysis.

It is important to acknowledge that there is currently no consensus on what outcomes are primary in interventions for shelter-based abused women, and not even about whether re-abuse incidence is an appropriate measure for evaluating IPV interventions. This is reflected in the wide variety of outcome measures included in the studies in this meta-analysis. Because it has previously been found that shelter-based abused women report needs in several domains, including mental health, safety and psychological needs, we grouped the outcomes in this meta-analysis accordingly. We propose that future RCTs evaluating shelter interventions include fewer, but well-validated, standardized measures for each of these three domains (see also Ramsay et al.). With respect to the social domain, we suggest that in addition to perceived social support, it may also be important to include measures for access to resources (such as finding housing or a job). Indeed, being financial autonomous protects women and enables them to leave abusive relationships, at least in the more developed countries.

The results of our meta-analysis should be interpreted with caution. As with all meta-analyses, the strength of the overall evidence is limited by the number and quality of the studies included. The 10 identified studies were few in number and very diverse in terms of type, goals, content, duration and intensity of the interventions. Also, the quality of the included studies was suboptimal: four studies were rated as poor and six as fair. Quality ratings were based on the information provided in the articles, but relevant information about the content of the interventions, fidelity assessments, the reporting of missing data and loss to follow-up was often lacking. In addition, due to the small number of studies, subgroups analyses could only be conducted for mental health outcomes, but not for abuse and social outcomes. Finally,
because only a few studies included follow-up data, we did not consider long-term effects of the interventions.

Many questions remain. Because the evaluated interventions were highly diverse, it is unclear which elements of them might have contributed to the positive effects. More research is needed that compares brief with intensive interventions, for example, or that compares different interventions with one another. For example, although it is often assumed that re-abuse is curbed by improving social support, this has not yet been systematically tested. It might also be important to conduct mediation analysis to examine what factors may influence the outcomes of interventions, and to test whether outcomes mutually influence one another. Finally, cost-effectiveness of shelter interventions needs to be studied.

In sum, we conclude that interventions provided during and after women’s stays in shelters are effective in improving mental health outcomes, in reducing the incidence of re-abuse and in improving social outcomes. Given the extent and complexity of the problem of IPV around the world, well-constructed studies evaluating shelter interventions in diverse populations are urgently needed.

**Supplementary data**

Supplementary data are available at EURPUB online.

**Funding**

This work was supported by the ‘Academic Collaborative Centre for Shelters and Public Mental Health’.

Role of the Sponsor: The funding organizations gathered in the Academic Collaborative Centre had no input in the design of this study in the collection, analysis and interpretation of the data; or in the preparation, review or approval of the manuscript.

**Conflicts of interest**: None declared.

**Key points**

- Interventions provided to women and their family members during and after stay in a woman’s shelter are effective in improving mental health, reducing re-abuse and improving social outcomes.
- It is recommended to reach consensus on desirable outcomes for shelter interventions.
- Randomized controlled trials evaluating (cost-) effectiveness of shelter interventions in diverse populations and cultural contexts are needed.

**References**


34 Moracco KE, Cole TB. Preventing intimate partner violence. Screening is not enough. JAMA 2009;302:568–70.