Mindful
Continuing Education

Intimate Partner Violence: Prevalence, Impact, and Treatment

Domestic violence and mental health: a cross-sectional survey of women seeking help from domestic violence support services

Background: Domestic violence and abuse (DVA) are associated with increased risk of mental illness, but we know little about the mental health of female DVA survivors seeking support from domestic violence services.
Objective: Our goal was to characterise the demography and mental health of women who access specialist DVA services in the United Kingdom and to investigate associations between severity of abuse and measures of mental health and health state utility, accounting for important confounders and moderators.
Design: Baseline data on 260 women enrolled in a randomized controlled trial of a psychological intervention for DVA survivors were analysed. We report the prevalence of and associations between mental health status and severity of abuse at the time of recruitment. We used logistic and normal regression models for binary and continuous outcomes, respectively. The following mental health measures were used: Clinical Outcomes in Routine Evaluation – Outcome Measure (CORE-OM), Patient Health Questionnaire, Generalised Anxiety Disorder Assessment, and the Posttraumatic Diagnostic Scale to measure posttraumatic stress disorder (PTSD). The Composite Abuse Scale (CAS) measured abuse.
Results: Exposure to DVA was high, with a mean CAS score of 56 (SD 34). The mean CORE-OM score was 18 (SD 8) with 76% above the clinical threshold (95% confidence interval: 70–81%). Depression and anxiety levels were high, with means close to clinical thresholds, and more than three-quarters of respondents recorded PTSD scores above the clinical threshold. Symptoms of mental illness increased stepwise with increasing severity of DVA.
Conclusions: Women DVA survivors who seek support from DVA services have recently experienced high levels of abuse, depression, anxiety, and especially PTSD. Clinicians need to be aware that patients presenting with mental health conditions or symptoms of depression or anxiety may be experiencing or have experienced DVA. The high psychological morbidity in this population means that trauma-informed psychological support is needed for survivors who seek support from DVA services.

Keywords: *domestic violence and abuse*; *intimate partner violence*; *mental health*; *posttraumatic stress disorder*; *anxiety*; *CORE-OM*; *depression*; *women*; *advocacy*

4% have experienced it in the past year (2). IPV is associated with depression, anxiety, posttraumatic stress disorder (PTSD), and substance abuse in the general population (3–5) and among women consulting in primary care (6, 7). There is evidence for a bidirectional effect (i.e. that women experiencing abuse are at greater risk of mental health conditions and that having a mental health condition makes one more vulnerable to abuse) particularly for depression, although there is a shortage of longitudinal studies to partition the directions of this effect (3). Qualitative research with survivors of IPV highlights the impact of abuse on the development of mental health problems (8). The few studies that have investigated the association between severity of exposure to IPV with mental and physical health problems reported positive associations (9–11). In these studies, the strength of association differed by type of abuse (9–12). Furthermore, Hegarty et al. (9) found that severe abuse is consistently associated with worse social coping, as well as increased levels of anxiety and posttraumatic stress symptoms. Abuse is also associated with poor self-reported physical health and pain, injuries, gynaecological and obstetric conditions, and difficulties carrying out daily activities (5, 13). Severity and type of PTSD (14) are also predicted by exposure to childhood abuse or maternal IPV (15).

Moreover, women who have recently experienced severe episodes of violence generally experience high levels of distress (9). Female survivors of IPV who seek advocacy support report high levels of abuse and depression when they first contact services (16, 17), higher than the general population (18). These levels decrease in time, independently of whether women are offered treatment or not (19, 20), and depression rates in women who have left a violent relationship up to 1 year earlier are similar to those in the general population (4).

Age may be a confounding factor in the relationship between exposure to IPV and mental health. Although younger women are at greatest risk of current abuse, older women have a greater lifetime experience; both current and lifetime experience increase the risk of mental health problems. Higher education and employment status are probably protective factors against IPV exposure (21–23). Socio-economic status, as well as recency and duration of abuse, therefore needs to be included in any analysis of the relationship between IPV exposure and mental health.

In this study, we aim to 1) characterise the demography and mental health of women who access specialist domestic violence and abuse (DVA) services in England and Wales; 2) investigate associations between the severity of abuse and measures of mental and physical health and quality of life, taking into account important potential confounders such as age and socioeconomic status, as well as important potential moderators such as exposure to direct maltreatment as a child (7, 21, 24) and prior history of mental health problems (3, 4).

Methods

Study setting and design

This study uses data from a cross-section of 260 women seeking help from two DVA services in the voluntary (non-statutory) sector in two UK cities, Bristol and Cardiff. Study participants were women recruited to the PATH (psychological advocacy towards healing) randomised controlled trial, testing the effectiveness and cost-effectiveness of a novel psychological intervention for survivors of DVA. Treatment was delivered by advocates or support workers called *specialist psychological advocates* in view of the specialisation they gained through the PATH training. Here we present findings from the baseline data we collected at recruitment. Sample size was determined by the need to detect reliable change in the main outcomes of the PATH trial (25). In this paper, the precision of the analysis is indicated by the confidence intervals of the estimated prevalence and associations.

Eligible participants were women 16 years or older who were experiencing DVA, which led them to seek support from a DVA agency in Bristol or Cardiff between 11 April 2011 and 4 June 2013. This included women who had experienced IPV or abuse (psychological, physical, sexual, or financial) from adult family members. Their first point of contact with the agencies, a support worker, screened them for other exclusion criteria: 1) psychotic illness; 2) severe drug or alcohol problem; 3) inability to read English; 4) current counselling, cognitive behavioural therapy, or other psychological treatments either in primary care or specialist psychiatric services.

Eligible women willing to discuss participation in the study were then contacted by a researcher who sought consent. At that meeting, women who consented to participation self-completed the baseline questionnaire on which this paper is based.

Data collection

The PATH baseline questionnaire contained validated measures of mental health and exposure to abuse from an intimate partner, a member of the woman's family, or another adult. It also contained questions on socio-economic variables including age, parity, and employment status; substance use and general health variables; and measures of childhood exposure to abuse and maltreatment (24). A researcher was present in the room when the women filled in the questionnaire, to provide assistance if requested.

Measurement

We used six scales to measure mental health (see Supplementary file). Symptoms of psychological distress were captured with the Clinical Outcomes in Routine Evaluation – Outcome Measure (CORE-OM), which measures symptoms of psychological distress in four domains: subjective well-being, problems and symptoms, functioning,

and risk to self or others (26). CORE-OM is a standard screening measure in counselling services across the United Kingdom (26), and there are normative values from general and clinical populations in the United Kingdom. We used the continuous clinical CORE-OM score, with values between 0 and 40 (26).

We measured symptoms of depression with the nine-item version of the Patient Health Questionnaire (PHQ-9). The PHQ-9 is routinely used in general practice in the United Kingdom to screen for symptoms of depression, and there are normative values for both clinical and general populations (27). We computed an indicator equal to 1 if the PHQ-9 score was greater than 9, that is, suggestive of major depression (28). Symptoms of anxiety were measured with the seven-item Generalised Anxiety Disorder questionnaire (GAD-7) (29). We computed an indicator equal to 1 if the GAD-7 score was greater than 9. We measured posttraumatic stress with Foa's Posttraumatic Diagnostic Scale (30), and adopted the threshold recommended for this population (at least 17 points) for our analysis on the binary outcome (14). The EuroQol EQ-5D-5L (31) measured health state utility on a scale from less than 0 (worse than dead) to 1 (perfect health). Finally, we measured quality of life with the SF-12 (acute form), a measure of health status. Specifically, we computed the SF-12 aggregate mental and physical health sub-scales, which capture respondents' physical and emotional health state and indicate whether these interfere with their daily lives and activities (32).

The measure of DVA was the Composite Abuse Scale (CAS). The CAS is a 30-item self-reported measure capturing emotional, physical, and severe abuse, as well as harassment (33). For our analysis we used a continuous version of the score, which can range between 0 and 150 (see Supplementary file). We preferred the continuous score to the binary (cut-off score: CAS ≥ 3) because of the high IPV exposure in our sample.

Recency of exposure was summarised by an ordinal variable that assigned higher values to more recent events. It varies between 0 (more than 12 months ago) and 4 (in the past month). Length of exposure varies between 1 (one occasion only) and 6 (for more than 5 years), increasing in the length of exposure. We summarised childhood abuse with a variable equal to 1 if the respondent had been the victim of either physical or sexual abuse in childhood. We also included a binary variable that denoted exposure to domestic abuse from a family member who was not an intimate partner, in order to account for exposure to multiple forms of abuse. Past mental health issues were self-reported by the women: the questionnaire asked whether they had experienced mental health problems such as depression or anxiety in the past. We coded all positive responses to this question as 1, and attributed a 0 score to all women who reported no problems. We used binary variables to capture whether the women had children younger than 4 years of age living with them and whether they were in a relationship. The indicator for cannabis use was set to 1 if the woman had used cannabis in the previous 12 months. We measured alcohol consumption with the AUDIT-C (Alcohol Use Disorders Identification Test – Consumption) instrument. We used a cut-off point of 3, which is thought to perform better for women and detects hazardous drinking (34). The women's age was measured in years, and their educational attainment with a categorical measure varying between 0 (no education) and 5 (bachelor's degree or higher). Their employment status was measured with a binary variable equal to 1 if the interviewee was not in work, that is, either unemployed, a student, or a retiree.

Analysis

The data from the questionnaire were entered into an Access database. The CORE-OM and PHQ-9, together with the urban centre and type of service variables, were entered twice independently to ensure accuracy. Consistency and logical checks were performed in Access.

All analyses were conducted in Stata 12.1 (35). We characterised the sample with descriptive statistics of all variables.

For continuous variables, coefficients and 95% confidence intervals were calculated with normal regressions. For binary variables, odds ratios and 95% confidence intervals were calculated with logistic regressions. We report the univariable odds ratios (coefficients) with 95% confidence intervals for associations between mental health and exposure to abuse. The odds ratio (coefficient) and 95% confidence intervals of the adjusted estimates accounted for age, education, employment status, relationship status, the presence of children younger than 4 years of age, alcohol and drug use, and help-seeking for mental health in the past (36). We also adjusted for non-IPV domestic abuse and childhood abuse, as well as recency and duration of exposure. To investigate whether recency, duration, or child maltreatment modified the association between exposure and mental health, we also tested for multiplicative effects (data available upon request). All adjusted estimates also account for site (Bristol, Cardiff) and service type (refuge, outreach services) to reflect stratification in the sample (25). We present a complete case analysis, so that all women who had not reported a value for one of the variables in the model were excluded from the analysis. The number of respondents used to compute the statistics is always reported. We also excluded from analysis the seven women (out of 251) who reported experiencing DVA only from other family members and not from intimate partners.

Ethical considerations

The study was approved by the South West National Research Ethics Service with specific approvals being

received from appropriate local research ethics committees. Informed consent was sought from each woman during the first meeting, before she filled in the questionnaire, and the research assistant offered support in case of distress while the questionnaire was being completed.

Results

The participating DVA services reported a total of 1,940 women requesting support during the recruitment period. We screened 66% of these women and 1,096 (86%) were eligible. Of these, 792 (72%) were approached and 263 (33%) recruited into the study. Three withdrew and 260 completed the baseline questionnaire, 13% of the women who originally requested support (Table 1). Language barriers and being in receipt of a psychological treatment accounted for 81% of ineligible cases (9% of initial throughput); time commitment represented the most common single reason why women declined recruitment after having been offered inclusion in the study.

For 26 of the 28 variables used in this analysis less than 10% of values are missing. The variable with the highest percentage of missing values is income (40%). In this paper we present the complete case analysis, and therefore we exclude income from the variables in our model, as we have two other measures of socio-economic status: level of education and employment. The women in our sample were 33 years old on average (Table 2); the majority had gained a City & Guilds diploma; almost 80% were not in formal employment.

Almost 70% of the women reported severe abuse, with an overall average of 57 on the continuous CAS measure (Table 3). Abuse episodes were relatively recent and had been sustained over time for the majority of women. Out of 251 women, 7 reported being victims of domestic abuse

Table 1. Recruitment

	Cardiff					Bristol				Total	
	Women's centre	Community outreach	Residential	Total	% of entered	Community outreach	Residential	Total	% of entered	*N*	% of entered
Entered service	**444**	**534**	**317**	**1295**		**519**	**126**	**645**		**1940**	
Screened	**162**	**408**	**209**	**779**	**60**	**372**	**121**	**493**	**76**	**1272**	**66**
Ineligible	*31*	*73*	*20*	*124*	*10*	*49*	*30*	*79*	*12*	*203*	*10*
Drugs and alcohol	6	9	1	16		7	0	7		23	1
Language barrier	14	14	7	35		20	27	47		82	4
Male	1	0	0	1		0	0	0		1	0
Psychotic	1	5	1	7		7	0	7		14	1
Psychological therapy	9	46	11	66		15	3	18		84	4
Eligible but not approached	*115*	*51*	*52*	*218*	*17*	*67*	*19*	*86*	*13*	*304*	*16*
One-off contact	29	0	0	29						29	1
SPA capacity	60	53	42	155		50	17	67		222	11
Researcher capacity	22	8	6	36		14	1	15		51	3
Other	4	19	4	27		3	1	4		31	2
Unable to contact/ declined	**4**	**59**	**58**	**121**	**9**	**65**	**10**	**75**	**12**	**196**	**10**
Approached	**16**	**284**	**137**	**437**	**34**	**274**	**81**	**355**	**55**	**792**	**41**
Did not consent to contact	*6*	*117*	*15*	*138*	*11*	*115*	*26*	*141*	*22*	*279*	*14*
Consented to contact	*10*	*167*	*122*	*299*	*23*	*159*	*55*	*214*	*33*	*513*	*26*
Met with researcher	**6**	**108**	**64**	**178**	**14**	**92**	**45**	**137**	**21**	**315**	**16**
Recruited	*4*	*95*	*47*	*146*	*11*	*86*	*31*	*117*	*18*	*263*	*14*
Not recruited	*2*	*13*	*17*	*32*	*2*	*6*	*6*	*20*	*3*	*52*	*3*
Wanted counselling	1	2	3	6		0	–	0		6	0
Time commitment	1	6	5	12		2	2	4		16	1
Other	0	5	9	14		4	4	16		30	2
Withdrawal	–	–	–	*0*		*2*	*1*	*3*		*3*	*0*
Total	**4**	**95**	**47**	**146**	**11**	**84**	**30**	**114**	**18**	**260**	**13**

SPA, specialist psychological advocates.

Table 2. Sociodemographic profile of the sample

	Mean	Median	%	Minimum	Max	Standard deviation (IQR)	*N*
Age	33	31		18	63	17	248
Maximum education level		City & Guilds and similar		None	Bachelor's degree or higher	(GCSE to A-level)	233
Income bracket		Up to £10,999		Up to £10,999	More than £60,000	(Up to £10,999, to £11,000–£20,999)	156
White			87			34%	253
Currently in a relationship			20			40%	250
Perpetrator is current partner			23			42%	236
Is a parent			81			39%	254
Has children under 4 years of age			37			48%	260
Works in the household			38			49%	237
Not in formal employment (excluding retirees and students)			78			42%	236
Hazardous drinking (AUDIT-C ≥ 3)			54			50%	251
Smoked cannabis in past 12 months			26			44%	245
Witnessed DVA as a child			52			50%	257
Was abused as a child			50			50%	257
Had a mental health problem in the past			82			38%	251

IQR, interquartile range; DVA, domestic violence and abuse.

from another member of the family and not from an intimate partner (2.8%; 95% confidence interval: 1.0%, 4.5%).

Two-thirds of the women reported clinical levels of psychological distress, with the total CORE-OM averaging 18 points (standard deviation: 7). At least 40% of women reported clinical levels of distress in all subareas of the CORE-OM, and at least 70% reported depression or anxiety symptoms (Table 4). Of 256 women, 197 (77%; 95% confidence interval: 71.2 to 82.9%) scored at least 17 points on the PTSD measure, the optimal threshold to identify this disorder (14), and 211 out of 256 (82%, 95% confidence interval: 77.6 to 87.1%) scored at least 15 points, the cut-off point recommended by Sheeran and Zimmerman (2002, in (14)). The measure of health state utility records a value of 0.6 (standard deviation: 0.3). Women in the general UK population have average

Table 3. Exposure to abuse

	CAS measure							
	Mean	Median	%	SD	Minimum	Maximum	Interquartile range	*N*
Severe abuse	6	3		8	0	33		248
Emotional abuse	31	31		16	0	55		248
Physical abuse	13	11		10	0	35		248
Harassment	8	7		6	0	20		247
Total abuse	57	49		34	0	136		245
Severe abuse >1			69	46%				248
Emotional abuse >3			96	20%				248
Physical abuse >1			92	28%				248
Harassment >2			86	35%				247
Total abuse >3			97	18%				245
Type of abuse, ordinal measure		SCA			None	SCA	(Physical and other – SCA)	251
Recency		In the past 3 months			More than 1 year ago	Past month	Between 6 months and less than 1 month ago	243
Length of exposure		Up to 3 years			Never	More than 5 years	Between (up to) 1 to more than 5 years	244

CAS, Composite Abuse Scale; SCA, severe combined abuse.

Table 4. Mental health, health utility, and quality of life measures

	Mean	SD	Median	Minimum	Maximum	*N*
		CORE-OM				
Subjective well-being	24	8	25	3	40	
Percentage with mean $\geq$1.77		74%				259
Problems	22	10	23	0	40	
Percentage with mean $\geq$1.62		70%				259
Functioning	20	8	20	2	36	
Percentage with mean $\geq$1.3		80%				259
Risk	4	7	0	0	30	
Percentage with mean $\geq$0.31		41%				259
CORE-OM	18	7	19	2	35	
Percentage with mean $\geq$1.29		76%				259
		Depression, anxiety, stress				
Depression (PHQ-9)	14	7	14	0	27	
PHQ-9 score $>$9		72%				258
Anxiety (GAD-7)	13	6	14	0	21	
GAD-7 score $>$9		70%				255
Post-traumatic stress (PTSD test for civilians)	26	12	27	0	50	
PTSD score $\geq$17		77%				256
		Utility				
EQ-5D-5L	0.6	0.3	0.7	$-$0.2	1.0	249
		Quality of life				
SF-12 Aggregate physical health	48	12	51	19	68	236
SF-12 Aggregate mental health	31	14	30	6	62	236

CORE-OM, Clinical Outcomes in Routine Evaluation – Outcome Measure; PHQ-9, nine-item Patient Health Questionnaire; GAD-7, seven-item Generalised Anxiety Disorder questionnaire; PTSD, posttraumatic stress disorder.

EQ-5D values between 0.81 and 0.94 in the age groups below 64, and never lower than 0.71 in older women (37). Finally, quality of life measures suggest somewhat worse mental and physical health states compared to the general US population (32).

The crude associations of severity of exposure to abuse with mental health distress and trauma are strong (correlation coefficient: 0.3 and 0.4 respectively, $p<0.0001$ in both cases), as is that with health state utility (-0.3, $p<0.0001$). Women who reported symptoms of depression reported an average abuse score of 61 (standard deviation: 33), compared to an average of 43 (standard deviation: 30) for women who did not report depression symptoms. Similarly, women who reported symptoms of anxiety recorded an average exposure score of 61 (standard deviation: 34), compared to an average of 46 (standard deviation: 30) for women with no reported symptoms of anxiety. The remainder of this section reports results from linear and logistic regressions of mental health states on exposure to abuse, controlling for modifiers and sociodemographic characteristics.

Table 5 shows positive associations between exposure to abuse and psychological distress and negative associations between health state utility and quality of life and abuse, all measured with good levels of precision, except for the mental health subcomponent of the SF-12 and the measure of depression, once we adjusted for confounders.

The severity of psychological distress increased with the severity and extent of abuse: for every additional point in the abuse score, women reported an increase of 0.081 points in the psychological distress score ($p=0.004$). Controlling for moderators such as childhood abuse, which increases the likelihood of exposure to abuse in adulthood (38), and sociodemographic characteristics slightly increased the size of this association, only slightly reducing the precision of the estimate.

The unadjusted association between exposure to abuse and posttraumatic stress was positive, with the measure of PTSD increasing 0.2 of a point for every unit increase in the measure of exposure to abuse ($p=0.004$). The size of this association was unchanged when we controlled for moderators and demographic characteristics.

Both measures of health state utility decreased as severity to exposure increased, with good precision for the physical health subcomponent of the SF-12 ($p=0.008$); precision was reduced once sociodemographic confounders were accounted for.

Associations between increasing exposure to abuse and symptoms of anxiety were positive and precisely estimated (Table 6).

Table 5. Associations between mental health and health state utility and severity of exposure to violence

Variable	Coefficient	Adjusted coefficient
Measures of mental health		
CORE-OM	0.081	0.1
95% CI	(0.050, 0.10)	(0.043, 0.2)
p value	0.004	0.013
N	245	174
PTSD	0.2	0.2
95% CI	(0.1, 0.2)	(0.1, 0.2)
p value	0.004	0.002
N	243	172
Measures of health state utility		
EQ-5D	−0.0028	−0.0037
95% CI	(−0.0038, −0.0018)	(−0.0052, −0.0023)
p value	0.003	0.003
N	238	170
Quality of life		
Aggregate physical health (T score)	−0.080	−0.093
95% CI	(−0.12, −0.040)	(−0.17, −0.012)
p value	0.008	0.035
N	228	165
Aggregate mental health (T score)	−0.10	−0.12
95% CI	(−0.18, −0.026)	(−0.23, 0.015)
p value	0.023	0.036
N	228	165

The first column of results reports coefficients from a normal univariable regression of the mental health or utility variable (COREOM, PTSD, SF-6D, EQ-5D) on exposure to abuse as captured by a continuous measure of the Composite Abuse Scale (CAS); the second column reports coefficients from a regression of CORE-OM, EQ-5D, SF-6D, and PTSD on CAS, and sociodemographic confounders (age, number of live-in children under 4, maximum level of education, use of drugs and alcohol, and work status) as well as measures of recency and length of exposure, previous mental health issues, exposure to non-IPV domestic abuse, and exposure to child abuse.

Unadjusted odds ratios suggest a small positive association between exposure to abuse and depression (odds ratio 1.02; 95% confidence interval 1.01 to 1.03). Adjusting for confounders leaves the association unchanged, but reduces the precision of the estimate.

The association with anxiety and PTSD is more precisely estimated than the one with depression. The univariable associations between exposure and the measures of anxiety and posttraumatic stress are positive. Controlling for moderators and other socio-economic variables suggests that the odds of being anxious or suffering from posttraumatic stress increase by 3% for every additional point in the score of exposure to abuse (95% confidence intervals: 1.02 to 1.05 and 1.03 to 1.04, respectively).

Table 6. Associations between binary mental health states and severity of exposure to violence

Variable	Odds ratios	Adjusted odds ratios
PHQ-9 $\geq$ 10	1.02	1.03
95% CI	(1.01, 1.03)	(0.99, 1.05)
p value	0.002	0.113
N	244	174
GAD-7 $\geq$ 10	1.02	1.03
95% CI	(1.01, 1.02)	(1.01, 1.05)
p value	$<$0.0001	$<$0.0001
N	241	174
PTSD $\geq$ 17	1.03	1.03
95% CI	(1.02, 1.03)	(1.03, 1.04)
p value	$<$0.0001	$<$0.0001
N	243	172

The first column of results reports odds ratios from a univariable logistic regression of the mental health variable (PHQ-9, GAD-7, PTSD) on exposure to abuse as captured by a continuous measure of the Composite Abuse Scale (CAS); the second column reports adjusted odds ratios from a logistic regression of PHQ-9, GAD-7, and PTSD on CAS, and sociodemographic confounders (age, number of live-in children under 4, maximum level of education, use of drugs and alcohol, and work status) as well as measures of recency and length of exposure, previous mental health issues, exposure to non-IPV domestic abuse, and exposure to child abuse.

In our analyses, none of the tests for interactions between severity of abuse and recency, length of exposure, and child maltreatment were statistically significant (data available from authors).

Discussion

Half of the women in our sample of IPV survivors had been exposed to IPV for up to 3 years and had experienced the last episode in the 3 months prior to getting in touch with the services. Half had been abused as children and more than four in five had had a mental health problem in the past. More than three-quarters reported symptoms of PTSD at the time they filled in the questionnaire. This finding is consistent with Howard and colleagues' systematic review of epidemiological studies of diagnosed mental illness that reported the risk of PTSD as higher among women exposed to IPV than any other mental health condition. This is an important finding for clinicians, particularly generalists, who often miss the symptoms of PTSD in the context of domestic violence (39). Given the ubiquity and severity of PTSD resulting from IPV (40), health services need to develop and implement specific IPV trauma interventions for survivors.

The participants in our study had substantially more psychological distress, as measured by the CORE-OM,

than the general and clinical populations of women in the United Kingdom. Their average score was four times higher than women in the general population, whose mean value is 4.8, and similar to women seeking psychological therapies in primary and secondary care, whose mean is 18.6 (41). The proportion of women who presented symptoms of depression in our sample was twice as large as that of women in UK general practice (27); for symptoms of anxiety, this proportion was three times as large (29). This profile is consistent with previous findings on women who seek advocacy support in the United States (16, 17) and Hong Kong (18).

Also consistent with other studies, we found that increasing severity of IPV was associated with worse mental health (10, 11, 36), especially anxiety and PTSD, even after controlling for confounders. In our population, exposure to recent IPV has a stronger association with symptoms of mental health illness than other known predictors: exposure to child maltreatment (3, 21), heavy drinking (23), or drug abuse (42), as well as a history of poor mental health.

Presentation of symptoms of mental illness in generalist or psychiatric practice should be considered a potential indicator of past or current IPV, or possibly non-partner domestic violence. It should prompt questions about abuse, as recommended in the WHO guidelines on intimate partner and sexual violence: '[H]ealth-care providers should ask about exposure to intimate partner violence when assessing conditions that may be caused or complicated by intimate partner violence' (43) including symptoms of depression, anxiety, PTSD, sleep disorders, suicidality, or self-harm.

We found a very small negative association between increasing exposure to DVA and our health-related utility measures. One explanation for this may be that this measure is not appropriate for capturing the health and quality-of-life-related impacts of exposure to DVA in a highly traumatised population. For example, some of the domain-specific items on the EQ-5D, such as 'I have [slight/moderate/severe] problems washing or dressing myself' are not likely to be relevant to this population.

The strengths of our study include its focus on women seeking help for DVA, providing a basis for designing interventions for that group; its relatively precise estimates of the association between DVA severity and symptoms of mental illness; and the relatively low proportion of missing data, with the exception of income, which we replaced with education level and employment status to include socio-economic status in the analysis. These two variables are positively associated with income in the general population.

A limitation of our study is that the women in our sample are a minority of the women who presented at the participating DVA services and may differ from the women who were not eligible for the trial, were not approached, or declined to participate. In terms of the main findings of our study – the high proportion of survivors of IPV with symptoms of mental illness and the association of these symptoms with severity of violence – it is likely that the potential bias is in a conservative direction: women receiving psychological therapy or with psychotic symptoms (5% of women expressing interest in participation) were excluded. However, as potential participants were being offered psychological therapy in the context of the trial, it is likely that women with more psychological distress would be more likely to consent. A more general limitation is that our findings cannot be extrapolated to the whole population of women who have experienced DVA, as only a minority seeks help from DVA services.

Overall, our findings are consistent with other studies on the association between IPV and mental health problems.

The high mental health morbidity among women seeking support from DVA services highlights the need for effective, trauma-informed support services for this population. Equipping non-specialist support workers in advocacy agencies with psychological skills to support survivors of IPV may represent an important avenue for improving survivors' well-being (44). Furthermore, particularly in resource-poor settings, upskilling non-specialist and non-medical personnel to deliver psychosocial support to women survivors of DVA may help engage hard-to-reach populations in a sustainable service framework. Were such interventions effective, they would very likely be cost-effective at improving survivors' well-being, given the high cost of IPV to individuals, health services, and society as a whole (45).

Authors' contributions

GF performed statistical analysis, wrote the first draft of the article, incorporated comments and finalised the article. RAD devised the intervention, commented on drafts. JB commented on drafts and managed the study. EH provided figures for table 1 and commented on drafts. TJP supervised the statistical analysis and commented on drafts. LS provided figures for table 1. GSF conceived the paper, and contributed to drafts. All authors viewed and approved the final draft. GSF and GF made the final decision to submit the article.

Acknowledgements

We are very grateful to all the women clients of Next Link and Cardiff Women's Aid who participated in PATH, our trial steering committee (Simon Gilbody and Nicky Stanley), the PROVIDE administrator (Patricia Martens), our Data Monitoring and Ethics Committee (Harriet Macmillan, Jackie Barron, Sandra Eldridge, and Michael King), and our patient public involvement group members (Patricia Chipungu, Jackie Evans, Josie Hughes, Maninder Kaur, Maria Rees, and Hyat Robbins).

Conflict of interest and funding

The authors report no conflicts of interest. This report presents independent research commissioned by the National Institute for Health Research (NIHR) under its Programme Grants for Applied Research scheme (RP-PG-0108-10084). The views expressed in this publication are those of the authors and not necessarily those of the NHS, the NIHR, or the Department of Health.

Cognitive behavioural group therapy for male perpetrators of intimate partner violence: a systematic review

Abstract

Background: Violence against intimate partners is a worldwide public health problem. Cognitive behavioural therapy delivered in a group format is widely used for the treatment of men's violent behaviour towards their female partners. A Cochrane review about the effectiveness of this therapy from 2011 revealed a lack of controlled studies. Our aim is to update the current evidence on the effectiveness of cognitive behavioural group therapy on men's violent behaviour towards their female partner.

Methods: The Cochrane Library, the Campbell Collaboration Social, MEDLINE, PsychINFO, CINAHL, SCOPUS, Embase, Open Grey, Grey Literature Report, and Sociological Abstracts were searched for studies investigating the effectiveness of cognitive behavioural group therapy on intimate partner violence published in the period of January 1, 2010, to February 12, 2018. Manual searches were also performed to identify randomized and non-randomized controlled trials. Data extraction was done in duplicate. The primary outcome was the reduction in violent behaviour, and secondary outcomes were physical health, mental health, quality of life, emotion regulation, and substance use. Study quality was assessed with the Cochrane Collaboration's risk of bias tool and the Risk of Bias In Non-Randomized Studies of Interventions tool. A narrative summary was used to describe the review findings.

Results: We identified six new studies that met the inclusion criteria: four randomized controlled trials and two non-randomized trials. Three of the randomized controlled trials found a reduction in intimate partner violence after treatment. The fourth randomized trial found that a subsample of responding partners reported a reduction in violence but no changes in the men's self-reported violence after treatment. No effect could be detected in the two non-randomized studies. Analysis of risk of bias revealed mixed results, indicating both strengths and weaknesses.

Limitations: Only a limited amount of studies which scored as "low quality" were available.

Conclusions: There is still insufficient evidence to confirm that cognitive behavioural group therapy for perpetrators of intimate partner violence has a positive effect. Future research should focus on randomized controlled studies distinguishing between convicted and non-convicted populations where violent behaviour is the primary outcome.

Trial registration: CRD42016041493.

Keywords: Batterer, CBT, Cognitive therapy, Group therapy, Intimate partner violence, Mental health, Perpetrator, Randomized controlled trials, Systematic review

Background

Intimate partner violence is a violation of human dignity and rights and includes various forms of physical, sexual, and psychological abuse [1–4]. In contrast to other types of violent acts, violence by an intimate partner often reoccur within the relationship and can go on for years [3, 5], and recidivism rates of 21% [6] to 42% [7] are reported. Violence against women is a global public health problem and studies on intimate partner violence suggest that nearly one third of women experience physical or sexual violence from an intimate partner during their lifetime [8]. Furthermore, WHO [9] and others [10] estimated that as many as 38% of female homicides globally were committed by male partners, and the global life-time prevalence of physical and/or sexual violence by an intimate partner was 30%. In addition, 20–75% of women have reported experiencing emotional violence [11].

Cognitive Behavioural Therapy (CBT) is one of the most actively researched psychotherapies and has received consistent empirical support for a host of mental health problems and conditions [12, 13]. In the treatment of aggressive behaviour, CBT interventions are now a commonly used approach to help different populations to regulate anger and aggressive behaviour [14]. The main techniques used in CBT focus on establishing a therapeutic relationship, behavioural change strategies, cognitive restructuring, modification of core beliefs and sche mas, and the prevention of relapse and recurrence. Cognitive theory suggests that psychopathology is characterized by the activation of a conglomerate of related or contiguous dysfunctional beliefs, meanings, and memories that operate in coordination with affect, motivation, behaviour, and physiological responses [12]. Different psychopathological conditions are associated with specific biases that influence how an individual incorporates and responds to new information [12, 13].

CBT is commonly used to address dysfunctional anger and violent behaviour among intimate partners. Research on the effectiveness of such interventions has yielded mixed results [15, 16]. A systematic review in 2007 identified six studies ($N = 2343$) which consisted of a mix of convicted and non-convicted male participants [16]. One study ($N = 218$) compared feminist-cognitive-behavioural-group-therapy with process-psychodynamic group therapy [17]. The second study ($N = 64$) compared a 12-week CBT-based substance abuse and domestic violence group with a 12-week twelve-step facilitation group [18]. The results were inconclusive in each of the two studies. The other four studies compared CBT with no intervention (1771 participants in total) [19–22]. Only one of these showed a statistically significant effect in favour of CBT [22]. A meta-analysis showed that the relative risk for violence was 0.86 in favour of the intervention group with a confidence interval of 0.54–1.38. However, a combination of a low effect size and a wide confidence interval led to the conclusion that there was insufficient evidence concerning the effectiveness of CBT. A revision of this study in 2011 failed to identify new randomized controlled trials, precluding any new meta-analyses [23].

The primary aim of this systematic review is to examine new evidence for the effectiveness of group-based CBT on men's violent behaviour towards their female partners. Secondly, we also review whether cognitive behavioural group therapy (CBGT) affects changes in self-reported physical health, mental health, quality of life, emotional regulation, substance use, and socioeconomic outcome among perpetrators.

Methods

The systematic review was registered in the International Prospective Register of Systematic Reviews (PROSPERO), no: PROSPERO 2016:CRD42016041493, and conducted according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) standards [24, 25].

Eligibility criteria

1) Adult male participants aged 18 years or older who had a history of physical, psychological, or sexual violence towards their female intimate partners.
2) Participants voluntarily referred or convicted to treatment.
3) Studies examining the effect of cognitive behavioural group therapy.
4) The control group condition should be classified as applying no intervention, another intervention, or a waiting list.
5) The study should report on type, frequency and recurrence of physically, psychologically and/or sexually violent behaviour.
6) Eligible studies were required to be randomized or non-randomized controlled studies published in peer-reviewed journals during the publication period of January 1, 2010, to February 12, 2018.
7) The studies were written in English, Spanish, or Portuguese.
8) Studies examining perpetrators of human trafficking, child exposure to intimate partner violence, or dating violence among adolescents were excluded. Also, studies examining other forms of therapy than cognitive behavioural group therapy (i.e. couple's therapy, individual therapy) were excluded.

Search strategy

A systematic literature search was conducted with the assistance of a medical research librarian (S.A.P) on

various databases: the Cochrane Library, the Campbell Collaboration Social, MEDLINE, PsychINFO, CINAHL, SCOPUS, Embase, Open Grey, Grey Literature Report, and Sociological Abstracts. The queries involved a combination of thesaurus and free-text terms that were optimised to identify studies on intimate partner violence and cognitive therapy in the respective databases (see additional file 1), building on a search strategy described by Smedslund et al. [16]. The search was limited to the period of January 1, 2010, to February 12, 2018, in order to find studies published since the review by Smedslund et al. [23]. In addition to examining the reference lists of included studies, the Journal of Interpersonal Violence and Journal of Family Violence were searched by hand for the relevant period.

Data extraction

Two authors (M.B.N and M.L.L-C) independently screened the abstracts and titles of the retrieved references and assessed the full text of potentially eligible studies. Discrepancies were resolved by discussion with a third author (T.P). Two authors (M.B.N and M.L.L-C) extracted data from all included articles by following the Template for Intervention Description and Replication (TIDIeR) [26]. The items extracted and recorded were the study design, setting, sample characteristics like age, voluntarily or court-ordered to treatment, as well as outcomes, treatment fidelity and length of follow-up. Moreover, type of intervention, type of control condition, measurement tools, and timing of the outcome assessment. The predefined secondary outcomes were also recorded. We contacted authors for further information if needed. The final decisions on which studies that met the inclusion criteria were made after discussion among the review authors.

Quality assessment

To determine the validity of randomized trials, three authors (M.B.N, M.L.L-C & T.D) worked independently using the Cochrane Collaboration's risk of bias tool [27]. The same authors assessed the remaining studies using the Risk of Bias In Non-randomized Studies of Interventions (ROBINS-I) tool [28]. This process was followed by a discussion between all authors about the methodological quality of the included studies.

Results

Search results

The database searches yielded 4570 unique references (see Fig. 1, study flow diagram depicted from RevMan) [29]. Hand searching of the bibliographies of the systematic reviews and articles selected for the full text review revealed one additional study with potential relevance [30]. The full text of 16 articles was retrieved and reviewed in detail. One of these studies was excluded because it investigated the effect of individual therapy [31], while another was excluded because it investigated couples' therapy [32]. One was excluded because it did not measure violent behaviour but rather thoughts and aggressive feelings [33], and four additional studies were excluded because the main intervention was not group CBT [30, 34–36].

Characteristics of included studies

A total of six studies were finally included in the study following the screening process. Table 1 presents the characteristics of the four randomized controlled trials with 731 participants [37–40]. Table 2 presents the characteristics of the two non-randomized studies with 854 participants, one was a controlled retrospective cohort study [41], while the other was a quasi-experimental study [42]. The studies were conducted in Norway [39], United States [37, 38, 40], Sweden [41], and Spain [42] and published in English except for the study by Boira et al. [42], which was published in Spanish. The interventions described were carried out within special health services, a community setting serving victims and perpetrators of domestic violence, a prison or probation service setting, and a university setting.

The participants were recruited voluntarily or court--referred for treatment. Most of the participants in the studies were convicted of intimate partner violence. However, there were large notable differences concerning participant samples between the studies, ranging from 26 to 528 in the randomized controlled studies and between 62 and 792 in the non-randomized studies. The mean age of participants ranged from 34 to 40 years old.

The interventions used in the studies varied in content, length and how they were delivered. Palmstierna et al. [39] investigated the effect of cognitive behavioural group therapy (CBGT) delivered in a combination of three to four individual sessions followed by 15 two-hour group sessions. Alexander et al. [37] investigated the effect of 26 sessions of standard CBGT gender re-education. Murphy et al. [38] investigated the effect of 20 weekly 2-h sessions CBGT. Taft et al. [40] investigated the effect of 12 weekly 2-h sessions of trauma informed group intervention.

With regard to the non-randomised studies Haggård et al. [41] investigated the effect of an integrated domestic abuse program (IDAP) consisting of a minimum of 8 individual sessions and 27 two-hour group sessions, while Boira et al. [42] investigated the effect of a 20-session manualised CBGT-program.

In all studies, the group leaders were therapists trained on intervention with perpetrators of intimate partner violence (psychologists, doctoral students in clinical psychology, clinical psychology graduate student trainee,

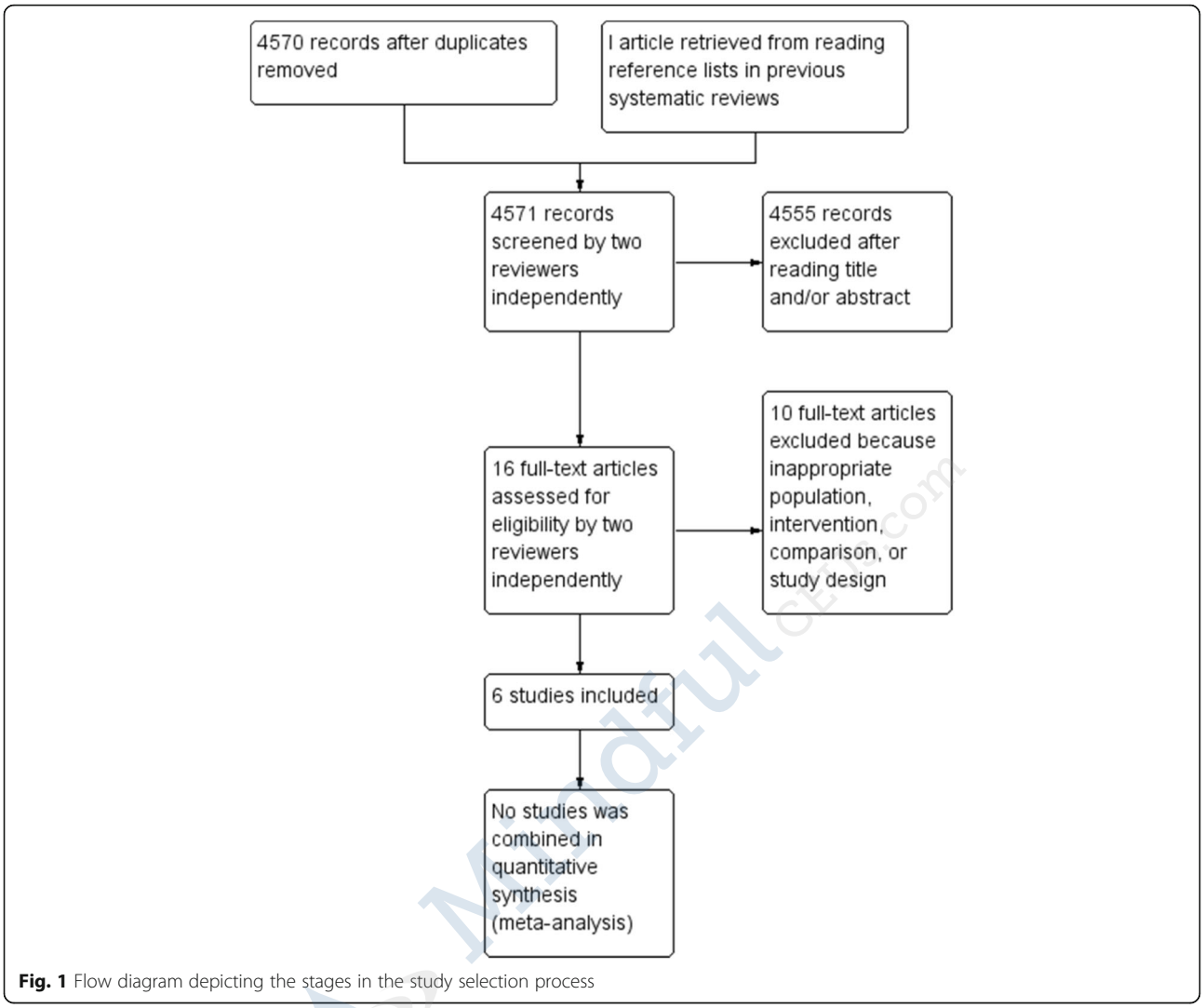

Fig. 1 Flow diagram depicting the stages in the study selection process

social workers, mental health nurses, or others with a university degree in behavioural science). The control groups were based on usual care [40, 41], an alternative intervention [37], a waiting list [39], or a comparison of the intervention with an open group format, individual therapy, or a waiting list [42], 20 sessions of standard individual cognitive therapy [38]. The intervention fidelity was measured in one study [37] by a blinded rater who listened to randomly selected audiotapes. Two studies reported treatment fidelity by recording group sessions followed by supervision to the instructors [38, 41].

Quality assessment

The risk-of-bias ratings for the randomized controlled trials are displayed in Fig. 2 a and b, depicted from RevMan [29]. All the included randomized trials in this review are judged as having poor quality. Additional file 2 shows more detailed information about the risk of bias ratings of each study.

Alexander et al. [37], Murphy et al. [38], Palmstierna et al. [39] and Taft et al. [40], score a high overall risk of bias in reporting according to the recommendations in the Cochrane Collaboration's tool for assessing risk of bias in randomised trials [27]. Alexander et al. [37] provide unclear information about the random sequence generation process, while the other three randomized controlled trials score a low risk of bias due to a detailed description of the random sequence generation. In all the four studies, the allocation concealment scores indicate that the risk is unclear due to inadequate description. Neither the participants nor the personnel were blinded to the treatment conditions in the four studies and therefore scored as high risk, although the research assistants making follow-up phone calls to the partners were blinded to the condition in the study by Alexander et al. [37].

With regard to incomplete data, two of the studies present no intention-to-treat analyses [37, 39]. While

Table 1 Characteristics of randomized controlled trials examining the effect of cognitive behaviour group therapy

Study, year, country	Setting	Population (N, mean age)	Intervention	Control condition	Outcome definition	Length of follow-up	Results: primary outcome
Alexander et al. [26], 2010, United States	Community setting serving victims and perpetrators of domestic violence	Male perpetrators (96.1% court-ordered) (*N* = 528, mean age 34.18 years)	Motivational Interviewing combined with Cognitive behavioural group therapy (SOCMI) 26 weeks	Cognitive behavioural group therapy (gender re-education), 26 weeks	CTS2, (psychological and physical aggression)	Perpetrator performed self-reports at 26 weeks' post-treatment. Partner assessments were performed at 6 and 12 months	No changes in self-reported violence. Significant reductions in partner reports of physical violence at follow-up in the SOCMI group
Murphy et al. [28], 2017, United States	A community-based domestic violence agency	Male perpetrators (N = 42, mean age 34.38 years)	Cognitive behavioural therapy, 20 individual sessions (ICBT)	Cognitive behavioural group therapy (CBGT), 20 weekly 2-h sessions	CTS2 (physical, psychological aggression, emotional abuse, relationship adjustment)	Perpetrator and partner performed self-reports at baseline and 3, 6, 9 and 12 months after baseline	CBGT produced equivalent or greater benefits than ICBT. Significant reductions in self-reported violence across conditions, with no between condition differences. Partner reports revealed more favourable outcomes for group treatment on measures of physical and psychological violence
Palmstierna et al. [25], 2012, Norway	Specialised outpatient mental health service	Male perpetrators voluntarily seeking therapy (*N* = 26, mean age 35.00 years)	Cognitive behavioural group therapy. 15 weeks 2 h sessions	Waiting list	CTS extended version (physical, material, any violence, verbal aggression)	Assessment after 15 weeks of treatment and after 15 weeks on waiting list as compared to baseline assessment	Significant reductions in self-reported violence in treatment group as compared to the waiting list group
Taft et al. [29], 2016, United States	Veteran Affairs hospitals Clinician-referrals, self-referrals, court-referrals	Male perpetrators; military veterans or service members (*N* = 135, mean age 37.85 years)	Cognitive behavioural group therapy, 12 weekly 2 h sessions (trauma-informed group intervention)	Treatment as usual	CTS2 (physical assault, psychological aggression) MINI, CAPS, MMEA	Perpetrators performed self-reports at baseline and 3 and 6 months after baseline. Partner assessments were performed at baseline and after 3 and 6 months	Significantly greater reductions in reported physical and psychologically intimate partner violence in the intervention group, self- and partner reports combined

CTS2 Conflict Tactics Scales–Revised, *CTS* Conflict Tactics Scales extended version, *MINI* The Mini-International Neuropsychiatric Interview, *CAPS* The Clinician-Administered PTSD Scale, *MMEA* Multidimensional Measure of Emotional Abuse. *CBGT* Cognitive Behaviour Group Therapy, *ICBT* Individual Cognitive Behaviour Therapy, *SOCMI* Stages-Of-Change Motivational Interviewing

Table 2 Characteristics of non-randomized studies examining the effect of cognitive behavior group therapy

Study, year, country	Setting	Population (N, mean age)	Intervention	Control condition	Outcome definition	Length of follow-up	Results: primary outcome
Haggård et al. [31], 2017, Sweden	Prison and probation offices	Consecutive sample of male IPV perpetrators: ($N = 792$, mean age 39.55 years)	Manual-based group program for male perpetrators (IDAP), including a pro-feminist psychoedu cational approach	Concomitant IPV offender controls who did not enter IDAP	Any new convictions for any violent recidivism and IPV during the follow up time	From time of recruitment unto study (2004–2007) until March 2, 2011. Mean time at risk, 4.6 years	19% ($N = 65$) of IDAP participants and 19% ($N = 84$) controls recidivated in violence against a partner or former partner
Boira et a. [32], 2013, Spain	Setting unclear. Treatment delivered by psychologists specialized in intimate partner violence	Male perpetrators convicted for IPV and court ordered to treatment ($N = 62$, mean age 39.70 years)	Three treatment modalities: 1. Structured group 2. Unstructured group (open group format) 3. Individual therapy	Waiting list	Police reports on new intimate partner violence	18 months	6.4% of the participants across the interventions were reported to the police for new intimate partner violence

CBT Cognitive Behavioural Therapy, *IDAP* Integrated Domestic Abuse Program, *IPV* Intimate Partner Violence

Murphy et al. [38], Palmstierna et al. [39] and Taft et al. [40] describe the distribution of attrition across groups, Alexander et al. [37] do not (they refer to another publication based on the same study). All four studies score as having high risk of bias for this item.

The risk of bias due to selective reporting is mixed across the studies. We found no protocol information on ClinicalTrials.gov for either Palmstierna et al. [39] or Alexander et al. [37]. Alexander et al. [37] only report on subjects completing the intervention and score at unclear risk of bias. Regarding other bias in the study of Alexander et al. [37], there is no power calculation or description of the how data were analysed and we suspect low statistical power. Palmstierna et al. [39] presents the results from self-reports of outcomes using the CTS and the associated *p*-values. However, the study only reports per-protocol results and gives no estimates of differences in reduced violence between the groups, hence this study is at high risk of bias on this domain. Taft et al. [40] did not report on the pre-defined secondary outcomes as stated in the Clinical Trials register, and score at high risk of bias on this domain. Murphy et al. [38] report all expected primary outcomes in the pre-specified way stated in the Clinical Trials register, and hence score at low risk of bias. With regard to other bias in the study of Murphy et al. [38], the imbalanced lack of compliance with allocated interventions between groups could cause bias and low statistical power. Hence, the study is at high risk of bias on this domain.

With respect to other sources of bias, Palmstierna et al. [39] were funded through the authors' employment at St. Olav's University Hospital and the Norwegian University of Science and Technology. Alexander et al. [37] were supported by the National Institute of Justice Grant. Murphy et al. [38] was funded by a grant from the National Institutes of Health, and Taft et al. [40] was supported by grants from the Department of Veterans Affairs and Department of Defence and through the use of the facilities and resources of the Providence Veterans Affairs Medical Center.

One of the non-randomized studies [41] is judged as having an overall moderate risk of bias, while the second study as having an overall serious risk of bias [42] (Table 3) according to ROBINS-I [28]. With regard to bias due to confounding, the study by Haggård et al. [41] statistically controlled for baseline recidivism risk that might have confounded the association between treatment status and recidivism in violent behaviour. The study by Boira et al. [42] scores as having low risk of bias due to confounding since the participants were selected from the target population and the study controlled for possible baseline confounding. Furthermore, the four groups had comparable sociodemographic characteristics.

Both studies are judged as having low risk of bias in the selection of participants for the study since both include all participants eligible for the target trial. Both studies clearly define the intervention and control groups and score as low risk on bias in the classification of interventions. Haggård et al. [41] followed an intention-to-treat approach and are therefore judged as having low risk of bias due to deviations from intended interventions. Boira et al. [42] provided insufficient information about intention-to-treat analysis, adherence to the interventions

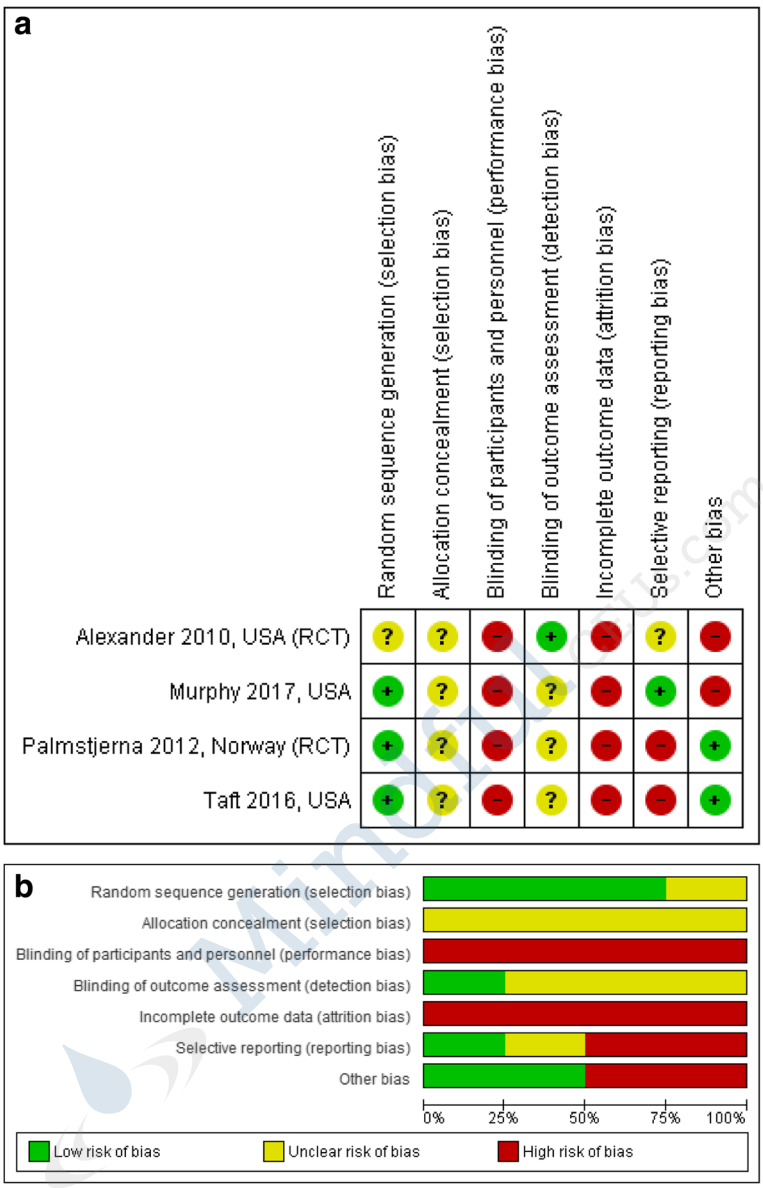

Fig. 2 **a** Risk of bias summary: review authors' judgements about each risk of bias item for each included study. **b** Risk of bias graph: review authors' judgements about each risk of bias item for each included study

and about the control group outcomes at post-test, hence there is no information to judge this item.

Risk of bias due to missing data is judged as low for Haggård et al. [41] since the study is retrospective and based on register data. Hence, no attrition from the study would affect the outcome. Furthermore, the study provides complete outcome measurements based on registry information. In the study of Boira et al. [42] there is insufficient information to judge this item. They report low attrition from the study but provide unclear information on missing data besides that.

The retrospective study by Haggård et al. [41] is judged as having low risk of bias in the measurement of outcomes since the results were already reported and the methods of outcome assessment were comparable across the intervention and control groups. Moreover, one assessor was blinded to recidivism data on any crime conviction in the past 5 years, any previous conviction of IPV, any previous conviction of a sexual offense, young age (below 21) at first known crime, any previous conviction of violation of a restraining order, current abuse or dependence on alcohol or drugs. The study by Boira et al. [42] is judged as having moderate risk of bias on this item due to a lack of blind outcome assessments and unclear information on the outcomes and intervention status for 18 months of follow up (the

Table 3 Risk of Bias summary according to ROBINS-I in the non-randomized studies

1st author, year, study design	Bias due to confounding	Bias in selection of participants into the study	Bias in classification of interventions	Bias due to deviations from intended interventions	Bias due to missing data	Bias in measurement of outcomes	Bias in selection of the reported results	Overall bias
Haggård, 2015, Sweden, Controlled cohort retrospective study [31]	Low risk Controlled statistically for baseline recidivism risk that might confound the association between treatment status and recidivism	Low risk All participants who would have been eligible for the target trial were included in the study	Low risk The intervention group (IG) and control group (CG) are clearly defined. Information about intervention status was obtained retrospectively	Low risk Followed an Intention-To-Treat (ITT)-approach. Other co-interventions that might have affected the outcome were balanced across IG and CG. The assessor extracting data was blinded to recidivism data	Low risk Attrition from treatment was described (IG 27%). The study had complete outcome measurement based on registry information	Low risk Retrospective study with already reported outcomes	Moderate Performed the study after the intervention was finished. There is no published protocol, making it difficult to know if the outcomes were pre-defined	Moderate
Boira,2013, Spain, Quasi-experimental study [2]	Low risk The participants were selected from the target population. The study had controlled for possible baseline confounding, and the three groups are comparable for sociodemographic characteristics	Low risk All participants who would have been eligible for the target trial were included in the study	Low risk The four groups are clearly defined	No information Insufficient information with respect to ITT-analysis and adherence to interventions	Moderate There is unclear information about recidivism data on the intervention groups at 18-months follow up (the outcome is presented as total participants, $N = 44$, making it impossible to separate the effects between the four groups status). Low attrition. It is unclear how missing data was analysed	Moderate Lack of blind outcome assessments. The methods of outcome assessment were comparable across three of the four groups for pre- and post-assessments. Errors of measurement occurs at 18-months follow-up (non-differential measurements are presented with respect to conviction)	Serious There is no published protocol, making it difficult to know if the outcomes were pre-defined The lack of differentiation between treatment modalities in presenting the results at 18-month follow-up makes it difficult to judge whether the observed effect is associated to group treatment	Serious

outcome is presented as the total participants ($N = 44$), making it impossible to separate the effects between the status of the four groups).

No published protocol was found for either of the studies [41, 42], making it difficult to determine whether the outcomes were predefined. Also, in the Boira [42] study a lack of differentiation between treatment modalities in presenting the results at 18-month follow-up makes it difficult to judge whether the observed effect is associated with group treatment. The treatment programs used by Boira et al. [42] (personal communication) were not compared, and they instead measured the effect of each program separately on new reports of intimate partner violence.

With respect to other sources of bias, Haggård et al. [41] reported indirect funding from the Swedish Prison and Probation Service through the authors' employment there. Boira et al. [42] did not report funding but had a collaboration agreement between the General Secretariat of Penitentiary Institutions and the College of Psychologists of Aragon.

Primary outcome: Effect on violent behaviour

The reported primary and secondary outcomes are summarized narratively given the considerable diversity of how they were assessed and the report of data in the studies included. Tables 1 and 2 display the primary outcome measures. Four randomized controlled trials [37–40] including 731 clients and 202 partners, and two non-randomized studies [41, 42] including 854 clients report outcomes on violent behaviour. Four studies [37–40] assessed violent behaviour using the Conflict Tactics Scale (CTS/CTS2) [43]. One study obtained register data from the Swedish Prison and Probation Service and court records on reconviction for violent crime against an intimate partner [41], while another study used register data on intimate partner violence reported to the police [42]. Only one study addressed sexual violence [41].

The small study by Palmstierna et al. [39] ($N = 26$) indicates a protective effect of CBGT on self-reported violence related to intimate partners as compared to the waitlist control, immediately after the intervention. This study also finds a significant correlation between low age and continued physically violent behaviour. The substantially larger study by Alexander et al. [37] ($N = 528$) find no differences with respect to perpetrator self-reports of violence at the end of treatment between men assigned to a group treatment program based on the stages of change model and motivational interviewing (SOCMI) and those in a program based on the Duluth model--inspired CBT. Of the 43% of partners who responded, fewer partners in the SOCMI group reported having experienced physical aggression at follow-up. Murphy et al. [38] ($N = 42$) find that cognitive behaviour group therapy produces outcomes equal to or better than individual cognitive behaviour therapy. The difference between the two conditions are statistically significant for partner reports of psychological violence and exceed a medium effect size for physical assaults and emotional abuse. Taft et al. [40] ($N = 135$) report that the intervention was more effective than the control condition in reducing psychological and physical intimate partner violence, with a small-to-medium between-group effect size.

Haggård et al. [41] report that 19% i.e. 65 of the 340 participants in the treatment group and 19% i.e. 84 of the 452 controls recidivated in violence against a partner or former partner during follow-up. In the small study by Boira et al. [42] ($N = 65$, four different comparison groups), the authors conclude that they cannot obtain any conclusive evidence on any of the many outcome measures. As for their primary outcome (police reports after 18 months on new intimate partner violence), they do not compare the three programs. Furthermore, they do not report differences between the programs and the control group, and 94% ($N = 44$) of the program participants (regardless of study condition) did not have any new incidents of intimate partner violent reported to the police at 18-month follow-up. The primary outcome for the control group is not reported.

Secondary outcomes

None of the studies included report treatment effects on physical health, quality of life, emotional regulation, and substance use after treatment. Only Boira et al. [42] report the effects of treatment on mental health as measured by the Symptom Checklist-90 (SCL-90). They found lower scores after structured group therapy in the SCL-90 depression dimension, Global Severity Index, and total symptom load.

Other measurements

Two studies [37, 42] assess the participants' readiness to change using The University of Rhode Island Change Assessment (URICA) [44]. However, the use of URICA by Boira et al. [42] is not satisfactorily explained. Alexander et al. [37] report a different outcome on physical violence: those with high initial readiness to change benefit more from group CBT than those with low initial readiness to change.

One study [42] assess empathy using the Spanish version of the Interpersonal Reactivity Index [45, 46], as well as hostility measured by the Spanish version of the Buss-Durkee Hostility Inventory [47, 48]. One study [37] assess risk factors for repeated violence using 12 items of the Danger Assessment Scale (DAS) [49]. Two studies [38, 40] use The Multidimensional Measure of Emotional Abuse (MMEA) [50] as an additional measure to assess

psychological intimate partner violence. One study [38] assess relationship adjustment using the Dyadic Adjustment Scale (DAS) [51], the participants' communication difficulties by partner reports on the Spouse Verbal Problem Checklist [52], and the participants' responses to challenging relationship scenarios using the Articulated Thoughts in Simulated Situations [53] paradigm.

Discussion

This systematic review evaluates and updates the evidence published on the effectiveness of cognitive behavioural group therapy for male perpetrators of intimate partner violence since the Cochrane review on this topic published in 2007 and replicated in 2011 [16, 23]. Only six studies met our inclusion criteria. Our main finding supports the results of the last updated review by Smedslund et al. [23] in 2011, that the evidence for this therapy is still inconclusive.

Three of the included studies found a reduction in physical violence among participants in the group-based interventions [38–40]. However, these studies were small and most of the findings relied solely on self-report from the perpetrators. The larger study by Alexander et al. [37] included 528 male participants and found only marginal differences in self-reported violence with respect to the type of treatment. Boira et al. [42] relied on police reports, which are known to capture only a small part of the actual incidents of intimate partner violence. Furthermore, the participants were not randomly selected for the different treatment modalities. Moreover, Boira et al. [42] had a wide range of outcomes without differentiation between primary and secondary outcomes. They reported only pre-post evaluations without comparing the group change differences. The study by Haggård et al. [41] reported the recurrence of intimate partner violence based on new convictions, which is also subject to the same limitation as the method by Boira et al. [42]: it does not necessarily show the true picture with regard to the violence that is actually occurring.

This review clearly confirms that self-reported outcomes like physical health, mental health, quality of life, emotional regulation and substance use are scarcely addressed when investigating the effectiveness of cognitive behavioural therapy for anger and aggressive behaviour. Future randomized controlled trials should therefore also address these outcomes.

A randomized controlled trial design is preferred when evaluating treatment effects due to confounding by indication. Nevertheless, only four of the six studies reviewed are randomized controlled trials. Furthermore, one of the studies has a limited sample size of only 26 participants [39].

When evaluating treatment effects, it is necessary to consider the treatment context [54, 55]. Delivering treatment within the prison service is different from an outpatient setting. The therapy is given in different settings across all the included studies, and most of the participants are involuntarily referred except for those examined in the study by Palmstierna et al. [39]. It is important to separate participants who are involuntarily assigned to treatment from those seeking treatment on their own initiative since they probably represent different subtypes of perpetrators with different associated risks of recurrent violence and treatment compliance [56, 57]. Earlier systematic reviews of cognitive therapy for perpetrators of intimate partner violence have not distinguished sufficiently between the different subtypes of perpetrators and the type of contexts where the treatment is delivered.

Limitations and implications for future research

Since only six studies met the inclusion criteria, the conclusions drawn from this review should be interpreted with caution. We found reasons to suspect that there is a high risk of bias across the included studies, mainly due to lack of blinding and incomplete outcome data reporting. Also, allocation concealment and other important domains were poorly reported and represents a threat to the certainty of the overall evidence.

Two of the studies used register data on convictions and police reports, and four studies used self- and/or partner reports. The lack of standardisation of the study design, follow-up time, and outcome measurement found in the included studies prevents us from performing a meaningful meta-analysis. Not reporting the outcomes according to the original random assignment violates the intention behind random assignment and makes the experiment less likely to take into account possible confounding by indication.

The scarce evidence on the effect of group-based CBT calls for well-conducted randomised controlled trials in different settings, as well as different and defined selections of perpetrators. The findings of this review underscore these important areas for future research, which is in line with earlier evidence on different treatment modalities for perpetrators of intimate partner violence [23, 58–60]. Our review and previous research on intimate partner violence programmes reveal that a combination of multiple theoretical models and treatment modalities are common in clinical practice, which makes outcome evaluations challenging [55, 61]. In future research, the elements of the treatment should at least be described clearly to make it possible to evaluate and compare treatment effects [54]. It is also important to ascertain the therapeutic adherence to the protocol, which will increase the attribution of effects or lack of effects to the intervention. Furthermore, non-randomized studies should publish protocols including a pre-analysis plan. It

is also recommended that randomized controlled trials use [26] and follow the CONSORT guidelines [62, 63].

Conclusion

The evidence is still inconclusive with regard to the effectiveness of group-based CBT in reducing violence from men towards their female partners – a situation that is due to a lack of high-quality randomized controlled trials on the subject. An important implication for future research in this area is to put an emphasis on describing the interventions in detail and reporting the study design and finally, how the study was carried out.

Our review also reveals that, so far, few studies have investigated how group-based CBT affects self-reported outcomes on physical health, mental health, quality of life, emotional regulation, substance use, and socioeconomic outcome among perpetrators. Based on our findings, future studies should adopt a randomized controlled study design with clear criteria for randomization, blinding and allocation concealment. Reduction of violent behaviour should be the primary outcome, as measured by both self-reports and partner-reports. A clear description of the investigated perpetrator population is warranted since studies of convicted perpetrators should be separated from studies of non-convicted.

Couple Therapy for Intimate Partner Violence: A Systematic Review and Meta-Analysis

Abstract

Intimate partner violence is a serious public health problem accompanied by substantial morbidity and mortality. Despite its documented impact on health, there is no widely recognized treatment of choice. Some studies indicate that couples suffering from situational violence may benefit from couples therapy, but professionals are cautious to risk the possibility of violent retaliation between partners. After a comprehensive literature search of 1733 citations, this systematic review and meta-analysis compiles the results of six studies to investigate the effectiveness of couple therapy as a treatment for violence. Preliminary data suggest that couples therapy is a viable treatment in select situations.

Intimate partner violence (IPV) is a prevalent problem that has serious adverse effects on human wellbeing. According to the National Violence Against Women Survey (NVAWS), each year about 4.7 million intimate partner victimizations occur among women in the United States (Black et al., 2010). Data from the same survey indicate that a quarter of women experience severe violence, with 50% suffering physical injury. The consequences of violence on victims' wellbeing range from acute physical injuries to long term mental and physical health consequences (CDC, 2003; Karakurt, Smith, & Whiting, 2013). Physical injuries due to violence could be mild, but are sometimes severe and even lethal (Campbell, 2002). Studies have indicated that in 2007 alone, over 1600 women died as a result of IPV victimization (Catalano, Smith, Snyder, & Rand, 2009). Furthermore, costs of IPV to society are extensive. A 2003 estimate calculated the total monetary cost of IPV against women in the United States to be close to $5.8 billion per year, on top of the unquantifiable toll it takes

on individuals' relationships, communities, quality of life, and well-being (NCIPC, 2003). This does not even include the costs associated with IPV against men.

Victims of IPV often exhibit comorbid conditions, including mental health issues like depression/anxiety and suicide attempts, as well as substance abuse, gastrointestinal disorders, and gynecological/pregnancy related issues (Black et al., 2010). There is a serious unmet need in the treatment of IPV, with many programs neglecting the subset of victims/perpetrators who wish to remain together. Many programs designed to prevent intimate partner violence solely target male offenders (Babcock & La Taillade, 2000) by providing gender-specific group therapy or individual treatments, while domestic violence shelters and other facilities traditionally support only female victims by offering therapy, support groups and educational programs. However, findings on batterer intervention programs indicate that these programs do not work as well as expected (Dobash, Cavanagh, & Lewis, 1996), with programs often experiencing high dropout rates and sometimes even having unwanted consequences (Babcock & LaTaillade, 2000; Mankowski, Haaken, & Silvergleid, 2002) such as the normalization of aggressive behaviors and antisocial peer influences (Murphy & Meis, 2008; Smith, 2007).

As an alternative approach to gender-specific therapy, however, controversy exists about whether couple therapy is appropriate, effective or even safe in treating potentially violent couples (Goldner, Penn, Sheinberg, & Walker, 1990). Clinical trials show that couple therapy functions on a systemic level (individual, couple, societal and intergenerational) and is effective when treating couples with dysfunctional relational patterns (Lam, Fals-Stewart, & Kelley, 2009). The relationship patterns which have shown improvement as a result of couple therapy have involved a myriad of dysfunctional qualities, including communication difficulties (Baucom, Sevier, Eldridge, Doss, & Christensen, 2011; Christensen, Atkins, Yi, Baucom, & George, 2006), conflict management issues (Davidson & Horvath, 1997), sexual problems (Clement & Schmidt, 1983; Dekker & Everaerd, 1983), and relationship complications (Cohen, O'Leary, & Foran, 2010), among other diverse concerns (Monson, Fredman, Macdonald, Pukay-Martin, Resick, & Schnurr, 2012).

Despite evidence supporting couple therapy, there has been little research on when this type of treatment is appropriate and advisable. It's possible that the key to effectively treating couple violence lies in accurately classifying the violence in order to facilitate the prescription of specifically tailored therapies. Gender-specific group therapy (Tolman & Edleson, 1995) is viewed as the standard treatment for IPV, even for couples experiencing situational violence (minor incidents initiated by both men and women), despite high drop-out rates (Babcock & LaTaillade, 2000), uncertain efficacy, and emphasis on inherent male fault. Clinicians, case workers and some researchers are concerned that talking about sensitive topics in couple therapy sessions results in increased tension, which in turn creates a risk of intimate partner violence for the victims as well as the worry of a violent retaliation after a session (Adams, 1988; Bograd, 1984; Saunders, 1986). As a result of these concerns, couples therapy is often deemed inappropriate for potentially violent couples and discouraged in the absence of solid findings. However, based on previous empirical evidence and theory there is reason to believe that couple therapy may provide an integral tool for treating situational violence among couples who do not wish to separate. Therefore, the aim

of this study is to conduct a systematic review and meta-analysis to determine the effectiveness of couple therapy in reducing violence in relationships.

Systematic Review and Meta-analysis Methodology

A systematic review and meta-analysis can be described as a method for systematically searching the literature and combining relevant study data from included studies to develop a single conclusion with greater statistical power. This is especially useful when analyzing interventions where there is lack of consensus, controversy, or small sample size (Higgins & Green, 2011). Meta-analysis is the process by which findings of existing studies are combined into an integrative statistical framework. This is particularly useful with interventions focused on IPV, where the numbers in each study are relatively small, limiting the power to detect differences when they do exist. We therefore conducted a systematic review and meta-analysis of couple therapy to better understand the effect of these interventions on IPV and violence recidivism. For this review, we were specifically interested in whether couple therapy can help some couples in reducing violence in their relationships. In this study, systematic review and meta-analysis was conducted by following the Cochrane Handbook for Systemic Reviews of Interventions guidelines (Higgins & Green, 2011). These guidelines describe the process of conducting high quality systematic reviews and meta-analyses in detail.

METHODS

Identification of Studies

A systematic review of the literature is conducted prior to conducting meta-analyses in order to reduce bias in the included studies. Researchers decide on the Populations, Interventions, Comparisons, Outcomes, Time and Settings (PICOTS) prior to the meta-analysis to prevent bias in the process. By identifying the PICOTS before finalizing the search strategy, the search of the literature is consistent across studies. We conducted a systematic review evaluating the effect of couples therapy on violence reduction for adult couples suffering from IPV. We developed inclusion and exclusion criteria using the PICOTS framework as follows: *P*opulation: Adult couples who are suffering from couple violence; *I*ntervention: Couple therapy; *C*omparison: Couple therapy vs individually oriented therapy/or no treatment control; *O*utcomes: Reduction in violence; *T*ime: Any follow-up period greater than 30 days; *S*etting: Outpatient.

We ran an electronic search in February of 2015 for any articles containing the keywords "violence" and some combination of "couples or couple or marital" with "counseling or therapy or treatment" in PubMed, Ebsco/Host (CINAHL Plus, PsycINFO, Humanities International Complete, and Women's Studies International), and Cochrane Library. Since the query systems categorize abuse as a form of violence and nest those results under the broader umbrella of violence results, searching for violence actually identified more studies than searching for abuse and other related terms would have identified. The team consulted a librarian before finalizing our search strategy. A hand search of the references of included articles and references from review articles identified during our electronic search was completed. This enhances the likelihood of identifying all relevant existing studies for

inclusion, and makes it possible for other researchers to replicate the search to acquire similar data to test reliability.

Two team members first reviewed each of the titles, then the abstracts and finally the full articles. The exclusion criteria were as follows for the title and abstract review phase: no original data, subjects under 18 years of age, no couple's violence, no couple therapy intervention, follow up less than 30 days after intervention ended, or no relevance to the key question. Discrepancies were decided by a third author. Finally, an article review was completed by two team members. For the full article review, exclusion criteria were expanded to the following: studies that did not use validated measures, articles that were not peer-reviewed (i.e. newspaper articles and dissertations), articles not in English, and articles with no comparison group (such as case studies). Conflicts were again decided through discussion and consensus among team members. We did perform a hand search of references from articles identified by our queries in the systematic review that did not meet our inclusion criteria (generally they were not RCTs) but that were relevant to our investigation to identify additional studies. Unfortunately, all of the studies that seemed promising from this hand search ended up being excluded due to our strict inclusion criteria. We did not search conference abstracts since conference findings often report preliminary results which may change with final publication. We did not hand search specific violence-focused journals since these journals typically reporting on the IPV studies are up to date in the electronic databases.

It should be noted that while there may be merits to including unpublished works like dissertations, theses, and conference presentations, these have not undergone the rigors of the peer review process, and we cannot confidently validate their methods and results within the scope of this project. Thus, in an effort to maintain a high level of quality control and ensure that all studies incorporated into this meta-analysis can stand up to the highest levels of scrutiny, we did not feel comfortable including these types of works at this time.

Data Collection

For each article, two of the team members abstracted the data independently using standardized data abstraction forms for study design, population, intervention, outcomes, and quality. For this study, outcomes that were extracted include violence recidivism, number of sessions, type of treatment, and settings in which the treatment occurred. For continuous outcome measures, such as those provided by the conflict tactics scale (Strauss, Hamby, Boney-McCoy & Sugarman, 1996) (commonly used scale for intimate partner violence research), the mean difference between groups and a measure of dispersion are extracted. If the between-group differences are not reported, the point estimate of the difference is calculated using the mean difference from baseline for each group. If the mean difference from baseline is not reported, available information is used to calculate this from the baseline and final values for each group (Higgins & Green, 2011). If there are no measures of dispersion for the mean difference from baseline for each group, the variance is calculated using the standard deviation of the baseline and final values, assuming a correlation between baseline and final values of 0.5. For dichotomous outcomes such as violence and no violence, the number and percent of events pre- and post- intervention are

abstracted. It is common in the literature for studies to report findings differently or occur in different populations, so heterogeneity is tested for to see whether or not the studies are homogenous enough to be pooled together. Data were then entered into MIX for meta-analysis software using Excel platform (Bax, Yu, Ikeda, Tsuruta, & Moons, 2006) and checked for any conflicts. Conflicts were discussed until a consensus was reached.

The quality of the studies included in the meta-analysis because sometimes results of meta-analysis are used for recommendations. There are several validated scales that are used to measure quality, depending on the type of study undergoing data abstraction. The Newcastle Ottawa Scale (Wells, Sea, O'Connell, Peterson, Welch, Losos, & Tugwell, 2010) is appropriate when measuring the quality of non-randomized studies, while observational studies' quality is often assessed with the Cochrane Risk of Bias tool (Armijo-Olivo, Stiles, Hagen, Biondo, & Cummings, 2012), which can be used for randomized and non-randomized studies. In this study, the quality of the studies included was measured in two ways. First, we completed the Cochrane Intervention Studies Scale (Armijo-Olivo, Stiles, Hagen, Biondo, & Cummings, 2012) for randomized controlled trials and the Newcastle Ottawa Scale (Wells et al., 2010) for observational studies. Second, two reviewers independently assessed the quantity of studies, study limitations, directness, consistency, precision, and publication bias across the studies using the GRADE criteria to understand any effects that the quality of evidence across studies may have had on our findings (GRADE Working Group, 2004).

We abstracted outcomes relating specifically to IPV changes among intervention and control groups. Changes in IPV were characterized by Male-to-Female violence frequency at pre-test and follow-up, recording mean and standard deviation values. In instances where both partners reported on the violence level, the higher values were used for analysis, as is consistent with general practice in this area of study. Violence frequency measures varied by study (though adapted from the Conflict Tactic Scale), using either the Modified Conflict Tactics Scale (mCTS), Revised Conflict Tactics Scale (CTS2), or Timeline Follow-Back Interview – Spousal Violence (TLFB-SV). All scales have been peer reviewed for validity.

Data Synthesis and Analysis

At the article review level, we identified a number of articles which reported on follow-up studies; these were aggregated with the original study. We created a set of detailed evidence tables. We conducted meta-analyses when there were sufficient data (from at least three studies) and studies were sufficiently homogenous with respect to key variables (population characteristics, study duration, and intervention characteristics). For studies having more than one arm, we chose the arm for inclusion that had the intervention most consistent with the other included studies in the meta-analysis. When more than one follow up interval was reported, we used the data from the follow up most similar to the other studies, in this case using a 12 month follow-up. Several of the studies involved multiple intervention groups. For our analysis, we used only one intervention group and one comparison group to reduce variability and to maintain independence of studies as required in a meta-analysis. For the experimental group, we prioritized selecting the interventions that involved individual couple therapy, followed by utilizing data from conjoint group couples therapy when

necessary. Since our primary goal is to examine the efficacy of couple therapy as a treatment, we chose to prioritize no-treatment controls as the comparison group, followed when necessary by using individual therapy data and then gender specific group therapy data. This was in an effort to minimize artifacts of comparing couples therapy to other forms of therapy. We evaluated heterogeneity among the studies considered for quantitative pooling with an I-squared statistic and considered an I-squared value > 50% to indicate high statistical heterogeneity (Higgins, Thompson, Deeks, & Altman, 2003). Meta-regressions are conducted when there are sufficient data and when studies are sufficiently homogenous with respect to key variables (population characteristics, study duration, and intervention type). The heterogeneity among the studies considered for quantitative pooling is tested using a standard chi-squared test, using a significance level of alpha less than or equal to 0.10. Heterogeneity among studies is also examined with an I-squared statistic, which describes the variability in effect estimates that is due to heterogeneity rather than random chance. Higgins, Thompson, Deeks and Altman (2003) consider a value greater than 50% to indicate substantial heterogeneity. The mean difference between groups is pooled using a random-effects model with the DerSimonian and Laird formula if there is substantial heterogeneity (DerSimonian, 1986). Stratified analysis or meta-regression is conducted if results indicate significant heterogeneity. We pooled the mean difference between groups in IPV using a random-effects model to account for any between-study heterogeneity (DerSimonian & Laird, 1986). Data analysis utilized MIX for meta-analysis software (Bax, Yu, Ikeda, Tsuruta, & Moons, 2006). Publication bias was assessed using funnel plots and Egger's test (Egger et al., 1997). Study weighting is the inverse variance method; therefore, it takes into account the standard deviation as well as the sample size.

Quality of Studies

The quality of evidence is the degree to which we can be assured that an estimate of the found effect is valid. The quality of the included studies is measured using validated quality scales. There are several validated scales that are used to measure quality, depending on the type of study undergoing data abstraction. The Newcastle Ottawa Scale (Wells, Sea, O'Connell, Peterson, Welch, Losos, & Tugwell, 2010) is appropriate when measuring the quality of non-randomized studies, while observational studies' quality is often assessed with the Cochrane Risk of Bias tool (Armijo-Olivo, Stiles, Hagen, Biondo, & Cummings, 2012), which can be used for randomized and non-randomized studies.

Strength of a Recommendation

The strength of a recommendation is the degree to which we can be assured that adherence to the recommendation will benefit rather than do harm (GRADE, 2004). In a systematic review and meta-analysis, we follow certain steps to improve judgements in the decision making process on which outcome is critical, and the overall quality of the decisions (randomized clinical trials are the gold standard). All of these judgments, recommendations, and the balance between harm and benefit depend on having a clearly defined question and considering potential outcomes that are likely to be affected. In this study, strength of a recommendation is measured using the GRADE criteria (GRADE, 2004).

Results

Identification of Studies

After boolean searching PubMed, Ebsco/Host, and Cochrane Library, we initially identified 1733 unique citations (Figure 1). We then systematically screened for relevance, first by title and then by abstract, which left us with 119 full text articles to examine for eligibility. Of these, 108 articles were excluded by at least two reviewers. This left 11 articles to be included in the systematic review. Since several articles were longer follow-ups of the same study, we ended up including a total of 6 studies for our quantitative synthesis.

Study Characteristics and Quality

The included studies were all randomized control trials conducted in person in the US without any pharmaceutical interventions. Four studies utilized individual couple therapy as the intervention, one study utilized conjoint group therapy, and one study used a combination of both. For the comparison groups, two studies used a no treatment control, two used gender specific individual therapy, and two used gender specific group therapy. (Table 1)

All of the studies measure the level of Intimate Partner Violence among couples, usually noting mean and SD of violence frequency. Two studies used the Revised Conflict Tactics Scale (CTS2), one used the Modified Conflict Tactic Scale (mCTS), and three used the Timeline Followback Interview – Spousal Violence (TLFB-SV). All of the studies measured Male-to-Female violence scores (two also recorded frequency of minor male violence and three recorded severe male violence frequency). Five studies reported Female-to-Male violence scores (two of which also recorded frequency of female minor and severe violence), and one study also reported total couple violence frequency. Additionally, two of the studies examined the efficacy of couple therapy for reducing IPV specifically among substance abusing populations.

Reporting of demographic data varied across the studies, but all of them provided details regarding average age and racial identification. In the 1998 study conducted by Schlee, Heyman, and O'Leary, the mean age of male partners was 38.4 years and the mean age of female partners was 36.24 years; almost all of the participating couples were Caucasian (~96%), with 2.7% identifying as African American. Of couples who participated in the Stith et al. study (2004), 63% were Caucasian and 25% were African American. Average ages for the male and female partners were 38.3 years and 35.6 years, respectively. In the study conducted by Fals-Stewart and colleagues in 2006, approximately 55% of couples described themselves as Caucasian and 33% described themselves as African American; mean age was 35.92 years for the male partners and 32.41 years for the female partners. The 2002/2009 study conducted by Fals-Stewart and colleagues involved approximately 69% Caucasian couples and 18% African American couples. Male partner mean age was 33.1 years and female partner mean age was 31.8 years. In 2009, Lam et al. conducted a study where the average ages of male and female partners were 34.4 years and 32.95 years, respectively; 63% of the couples identified as Caucasian, while 20% identified as African American. The Bradley et al. studies (2011, 2012, and 2014) involved male partners with an

average age of 35 years and female partners with an average age of 34 years. 83% of the couples identified themselves as Caucasian, and 14.5% identified themselves as African American (it should be noted that in this study, participants could select multiple racial descriptors). Overall, couples who participated in studies included in this meta-analysis had a mean age of 34 years, with approximately 70% of couples identifying themselves as Caucasian and approximately 21% of couples identifying themselves as African American. All of the studies involved heterosexual couples. None of the studies identified any demographic variables as having any correlation with outcome variables.

After two reviewers separately assessed the studies included in this meta-analysis, we determined that all were moderate to high quality (Figure 2). Attrition bias seemed to be the most problematic issue overall, as most of the researchers did not provide information about drop-out characteristics. It is unclear if this would have influenced reported results, and thus changed our effect sizes. We encourage future studies to provide as much information as possible about withdrawals in an effort to improve study quality and more accurately and completely represent findings.

Descriptive Analysis of Included Studies

We provide a brief description of each of the six included studies. In 1998, Schlee, Heyman, and O'Leary conducted a study on violent couples investigating the differences in outcomes between couples who participated in a conjoint group therapy program, Physical Aggressive Couples Treatment (PACT), versus couples who participated in gender specific group therapy. After analyzing the data in 1999 (O'Leary, Neidig, and Heyman) and 2006 (Woodin and O'Leary), both interventions demonstrated violence reduction at follow-up, but the only significant differential effect of treatment type that was identified revealed more improvement on marital adjustment among husbands in conjoint treatment. Predictors of recidivism did not vary by treatment. The results of a 2004 study conducted by Stith, Rosen, McCollum and Thomsen showed benefits of conjoint group therapy over gender specific group therapy for male violence recidivism and aggression levels, but did not identify a similar effect among the participants of individual couple therapy (although both conjoint treatment programs showed lower recidivism rates than gender specific treatment, according to female partner reports at 2 year follow-up). In 2006, Fals-Stewart and colleagues examined the efficacy of couple therapy when treating alcoholic females, and found that participants in behavioral couples therapy (BCT) showed significantly greater improvement in dyadic adjustment than participants in either individual-based treatment (IBT) or PACT, with couples who went through BCT reporting higher dyadic adjustment and reduced partner violence at follow-up. Fals-Stewart and Clinton-Sherrod conducted a similar study in 2009 involving substance abusing men and their partners, comparing IBT with BCT. Participants in BCT reported significantly lower levels of IPV and substance use at follow-up compared to the IBT group, and treatment assignment appeared to be a significant moderator of the day-to-day relationship between IPV and substance use. In a 2009 pilot study investigating the effects of another form of treatment, Parent Skills with Behavioral Couples Therapy (PSBCT), Lam, Fals-Stewart, and Kelley compared PSBCT to BCT and IBT to assess its effect on parenting, relationship conflict, and substance abuse in family and dyad relationships. They found that BCT showed clinically meaningful effects over IBT for

treating substance use, reducing partner violence, and improving dyadic adjustment. In 2012, Bradley and Gottman implemented the Creating Healthy Relationships Program (CHRP), which is a couple and relationship education program designed to reduce IPV in low income situationally violent couples. Results showed that participation in CHRP is associated with higher rates in attitudes that reflected healthy relationship skills as well as a reduction in IPV for follow up. Meta-analysis of the study data described above has revealed a modest but significant positive effect resulting from the use of couple therapy as a treatment for IPV, as explained next.

Meta-Analysis of IPV Reduction

Our meta-analysis results indicate that intimate partner violence can be significantly reduced through the application of couple therapy when compared to an active comparator or no treatment control (weighted mean difference –0.84; 95% confidence interval of –1.37 to –.30) (Table 2, Figure 4). Further testing revealed that the studies included in the analysis have low heterogeneity (Figure 4), so are suited to pooling for accurate treatment comparison. Specifically, after evaluating the heterogeneity of the included studies, we determined a Q value of 2.38 (P = .79), with I^2 statistic equal to 0.00% (confidence interval 0.00% – 74.62%). Inter-trial variance value t^2 was also 0 (confidence interval 0 – 1.75), and the ratio of generalization of Cochran's heterogeneity statistic H equals 1 (confidence interval 1 – 1.99). Thus, overall heterogeneity between studies is deemed to be low. Our data were heavily influenced by a single study due to its large sample size and smaller inter-individual heterogeneity relative to the other studies. However, effect sizes of most of the other studies were similar to this one larger study. A sensitivity analysis where we remove the largest study shows similar effects but loses statistical significance. Given that all effect sizes are in the same direction, we anticipate that further work will likely solidify this early evidence. We decided to keep this study because it passed all of the exclusion criteria, it is relatively high quality, and we used only the follow up data.In short, evidence is classified into four categories: (1) "high" grade (indicating high confidence that the evidence reflects the true effect, and further research is unlikely to change our confidence in the estimate of the effect); (2) "moderate" grade (indicating moderate confidence that the evidence reflects the true effect, but further research could change our confidence in the estimate of the effect and may change the estimate); (3) "low" grade (indicating low confidence that the evidence reflects the true effect, and further research is likely to change our confidence in the estimate of the effect and is likely to change the estimate); and (4) "insufficient" grade (indicating evidence is unavailable or the body of evidence has unacceptable deficiencies, precluding reaching a conclusion). According to the GRADE criteria, we determined that the strength of evidence was moderate for this finding due to consistency of study findings, the relative high quality of included studies, the direct measurement of a clinically relevant outcome and precision of the results. After grading the strength of the evidence for couple therapy vs controls, we determined that risk of bias was medium, evidence was consistent and direct but imprecise, the magnitude of effect and SOE were moderate, and no publication bias was detected.

Discussion

Previous research has examined the feasibility and effectiveness of couple therapy based treatments for situational couple violence across several studies. In these studies, researchers have conducted couples therapy when working with couples experiencing relationship violence, and have reported differing but promising results. The aim of this study was to conduct a systematic review and meta-analysis to understand and synthesize data from the previous literature on the effectiveness of couple therapy. To our knowledge, this is the first study to do so. It is hypothesized that by aggregating data from previous studies, it may be possible to provide substantial evidence in support of using couples therapy as a treatment for IPV in certain circumstances. Our data support this view.

In this systematic review and meta-analysis, we abstracted data from six studies. This was a high quality systematic review, using only moderate to high quality studies. Although our effect sizes were only moderately significant, the analysis of the combined data from these studies provides evidence supporting the idea that couple therapy is a slightly better treatment approach than standard treatments when working with violent couples. This indicates that couple therapy can be an effective way to prevent intimate partner violence in certain situations. However, further research into what couple and relationship characteristics predict greater effectiveness of couple therapy is necessary before the widespread adoption of conjoint therapy as a standard form of treatment is really feasible.

According to Johnson and Leone's (2005) study using data from a large national survey, about 65% of violence in relationships exhibit the characteristics of situational couple violence. Importantly, some couples experiencing situational couple violence may prefer to continue their relationship with a desire to end the violence, enhance the quality of their marital relationship, and successfully parent their children (Stith & McCollum, Rosen, Locke & Goldberg, 2005). Gender specific treatments often operate in accordance with the Duluth Model (Pence & Paymar, 1993), which emphasizes the idea that the male perpetrator is psychologically driven to violent, oppressive, dominant behavior, and must be isolated from the victim. This is not conducive to treating couples who struggle to control their emotions but are committed to remaining together, and may actually deter couples from seeking help. Unfortunately, many social services available to female victims are also contingent on a commitment to leave the relationship. This means that many couples may be unable to access the counseling they need to improve their quality of life as a family, likely resulting in the progression of violence. For these reasons, research into couple therapy as an approved method of treatment for situational couple violence is critical.

While there are some other clinicians and researchers who argue that it is better to help these couples in a controlled environment with an experienced facilitator/therapist to guide them through their conflict issues together, rather than leave them on their own when they have already been aggressive towards each other. It is certain that safety precautions should be addressed to protect potential participants. There are a number of interventions cited in the literature for clinicians and researchers to help them address their safety concerns for partners during and in between therapy sessions (Karakurt et al., 2014). However, it is undeniable that certain issues associated with situational couple violence are more

effectively addressed with both partners present, so that the couple can grow stronger attachment bonds and support one another through the process.

Summary of Findings

This systematic review and meta-analysis on the effectiveness of couple therapy in violence recidivism indicates a positive impact of couple therapy. Results of the preliminary meta-analysis with pooled data from 6 studies with 470 participants indicate that couples therapy significantly reduces intimate partner violence by point estimate −.84, z =−3.07, ($p<.05$) with the confidence interval of −1.37 to −.30. Further research is needed to confirm these findings, but there is certainly reason to re-evaluate the role of couple therapy in IPV treatment and cautiously increase its application.

Sources of Bias

There are various possible sources of bias when conducting a meta-analysis. The magnitude of the bias and potential sources of bias are taken into account when conducting a meta-analysis, to provide more reliable and valid findings. These biases include selection bias, performance bias, attrition bias and reporting bias (Higgins & Green, 2011). A review of the studies analyzed in this meta-analysis revealed low risk of bias in all areas except attrition and publication (Figure 3). There does not seem to be any influence from selection, performance, or detection biases. The studies maintained consistent methods between experimental and control groups.

Selection bias is defined as differences in baseline characteristics of compared groups. To ensure any differences or similarities in participant characteristics between groups are not systematically similar to the point of potentially influencing study findings, interventions are randomly allocated. This is not always possible in clinical trials due to ethical considerations of withholding treatment. In meta-analysis, randomization is considered when rating quality of evidence to verify that studies followed a specific rule on how participants were allocated to different treatment options (Higgins & Green, 2011). Recruitment, screening, and group assignment procedures were the same between experimental and control groups for all six studies included in this analysis, so selection bias was judged to be low.

Performance bias is defined as systematic differences in how treatment/intervention is provided. Receiving higher quality of care or exposure to factors other than the intervention itself can influence findings. Effective blinding procedures are used to ensure participants receive similar amounts of attention, additional treatment, and diagnostic investigations. However, depending on the study or disease, blinding is not always possible (Higgins & Green, 2011). Detection bias is similar, but it results from systematic differences in how treatment outcomes or recidivism are measured. Studies in this meta-analysis appear to have given equal treatment attention to all groups regardless of intervention, minimizing performance bias, and outcomes were measured consistently between groups, so there was no evidence of detection bias.

Attrition bias is described as systematic differences between groups who complete the study. It is possible that participants with certain characteristics drop out from a study more frequently than participants with other characteristics, leading to attrition bias (Higgins &

Green, 2011). Given the complicated nature of situational couple violence and the discomfort that can arise from couple therapy as a result, it is not surprising that some of the six studies we analyzed suffered from attrition. Drop-out rates were sometimes high, and bias may be a concern because not all of the studies reported statistics and data for the withdrawal population (Figure 2).

Reporting bias, also referred to as "publication bias", is described as systematic differences among reported and unreported results. Generally, there is more published literature on positive findings, where an intervention works for a certain condition or situation, than on non-significant findings (Higgins & Green, 2011); this can inflate overall effect sizes found in meta-analyses. For the most part, individual studies in this meta-analysis did not show publication bias, in that they did not show evidence of withholding data. However, the meta-analysis itself does suffer from publication bias. To counteract the effect of the limited number of publications on negative (null) findings included in the meta-analysis (Rosenthal, 1979), unpublished data were sought out by sending individual e-mail requests to authors of articles reporting on violence where the quantitative data had not been reported. After contacting Dr. Harris regarding "*A Comparison of Treatments for Abusive Men and their Partners within a Family-Service Agency*" (1988), Dr. Markman regarding "*Preventing Marital Distress Through Communication and Conflict Management Training: A 4- and 5-Year Follow-Up*" (1993), Dr. Taft regarding ""*Strength at Home" Intervention to Prevent Conflict and Violence in Military Couples: Pilot Findings*" (2014), Dr. Dunford regarding "*The San Diego Navy Experiment: An Assessment of Interventions for Men Who Assault Their Wives*" (2000), and Dr. O'Farrell regarding "*A Randomized Clinical Trial of Behavioral Couples Therapy Versus Individually Based Treatment for Women With Alcohol Dependence*" (2014), no additional data were received.

Limitations

While the studies analyzed here demonstrate reasonably good quality, they do present several limiting factors. Perhaps the most obvious of these is the low number of studies we were able to incorporate in the analysis. There were very few study designs that actually met our screening criteria. Several studies appeared to have relevant data, but we were unable to use them because they either did not have a comparison group or they were using a matched-sample control, which made the results meaningless for this analysis because the matched samples are survey data of the general population matched only along demographic variables. Matched samples do not control for comparable substance abuse or IPV levels, introducing unacceptable levels of performance and detection bias. We also found that several studies did not report data consistently. Some articles did not publish standard deviation values, or did not report both pre and post data. One study did not show separate results for the intervention and control groups, reporting overall violence instead. Unfortunately, our attempts to contact the authors of these studies did not result in additional data. Although our strategy to screen out poor quality studies left us with very few articles to analyze, it did result in low heterogeneity and a more valid result. Still, without a sufficient pool of applicable literature, it is not possible to make significant comparisons between different treatment approaches, and there is greater risk of having results overly influenced by a single large study, as was the case with this present analysis. It is important to note that

there are some concerns regarding some of the Fals-Stewart study (Heisel, 2010). However, a substantive and methodological review conducted in 2010 supports the use of the behavioral couple therapy program developed by O'Farrell, Fals-Stewart and colleagues for the treatment of substance abuse, providing evidence in favor of the veracity of Fals-Stewart's results in this instance (Ruff, McComb, Coker & Sprenkle, 2010). Additionally, we performed a trim-and-fill analysis to see if removing any of the studies we included in our meta-analysis would alter the direction of our results. We found that the direction was unchanged for all of the studies, even the Fals-Stewart study, despite its obvious weight from having such a large N value (N = 207). Thus we decided to include the study in our meta-analysis.

We were able to collect some information on Female-to-Male violence as well as overall couple violence, but not all of the studies measured this so we did not include it in our analysis. Furthermore, while some studies reported information about probation and involvement of the legal system, other studies did not. Similarly, while some studies measured minor and severe levels of violence, we did not feel that combining the limited data available would significantly increase the power. As a result, we only analyzed pre and post Male-to-Female violence data.

For the most part, demographics were consistent across the studies we analyzed, but methods of reporting demographic data were not. It would be helpful if the field developed a standard for collecting and reporting data. Some studies use means to report demographics while others use ranges. There also does not seem to be a consensus on whether it is more appropriate to provide information on couples as a unit or information on partners separately, and some studies only indicate overall data without reporting separately by intervention and control group. This analysis may also be affected by the fact that many of the studies we have included are specifically investigating violence among substance abusing populations. While this does not prevent us from seeing if there is any correlation between violence levels and participation in couples therapy, it does limit our ability to generalize the observed effect across a wider population. Ultimately, this analysis is limited by the difficulty in synthesizing data from such a wide variety of study designs and reporting methods, as well as being limited by the scarcity of moderate to high quality studies that have been published on this topic.

Future Research

The limitations of meta-analysis are broadly known, as a meta-analysis is only as rigorous as the included studies. For this reason, we call for more rigorous and randomized empirical work on the effect of couple therapy. Extant research has barely begun to scratch the surface of the effectiveness of couple therapy for couple violence, and more research should explore how, when, and why this style of therapy may be most effective. Additionally, since many studies were excluded due to missing a comparison group or having no baseline violence data from the comparison group, we highly recommend all future studies include at least one control group, ideally multiple (e.g., a waitlist group and another therapy group). Additionally, future research should be mindful of the comorbid nature of IPV with other serious health concerns like mental and substance abuse disorders, and consider screening

for IPV in a variety of healthcare settings. Treatment should also take comorbid conditions into account, as it is currently unclear what the causational role is between IPV and co-occurring conditions, and the influence of undiagnosed mental and substance abuse disorders may drastically affect couple therapy efficacy.

Additionally, although there were not enough data collected for this meta-analysis to analyze the effect of couple therapy on minor vs severe forms of violence, or female perpetrated violence against male victims, we do feel that these are important avenues of research to pursue. While our data demonstrate a moderate effect of couple therapy in reducing male perpetrated situational relationship violence against females, this cannot be extended to all instances of couple violence. As professionals continue to discover the extent of previously underestimated female perpetration in situational couple violence, it will become increasingly necessary to understand whether couple therapy is an effective method for reducing that violence as well. Severity of violence and effectiveness of couple therapy should also be more rigorously investigated, as understanding the role of severity in recidivism may help clinicians to more easily determine treatment appropriateness and assess safety risks.

There are many standard treatments for IPV, but few involve couple therapy. While this meta-analysis speaks to the efficacy of couple therapy as a treatment for mild-moderate cases of IPV, it does not investigate the efficacy of other specific forms of intervention, nor does it imply that couple therapy is the best treatment. We have only determined that couple therapy works in certain situations in reducing violence recidivism. Future research should analyze the efficacy of other standard forms of treatment, so that policy makers and clinicians can use this research as a guide for determining the best treatment for each couple.

Conclusions

Policy & Clinical Implications

Overall quality of evidence was moderate, indicating that there are benefits of couple therapy for reducing violence in the relationship, but these benefits may be variable across couples. Our sample was not very diverse. We caution therapists to carefully consider their patients cultural background before determining whether couple therapy is an appropriate treatment. Our sample was also relatively young on average, which may limit the generalizability of our findings to older populations. IPV has been found to be more prevalent among younger adults, but this may speak to poor reporting or high levels of morbidity. While we believe that couple therapy is still an effective treatment among older adults, we encourage therapists to carefully assess each couple before prescribing any treatment, as the interplay of IPV and elder abuse is not clear without further research. Setting of treatment may be an important aspect as well, since couples may respond differently to therapy depending on the environment. We did not find any influence of setting on couple therapy efficacy, but more targeted research would be needed to determine if this is an important factor when treating IPV. We also wish to stress that our findings at this time can only be applied to instances of mild to moderate situational couple violence. To be clear, ‘situational couple violence’ refers to mutual mild violence among partners in response to specific stressors or life events as a means to resolve conflict, in contrast to

‘characterological couple violence’ or ‘intimate terrorism’, which are predominantly characterized by partner domination and severe physical abuse. Further research is needed to understand the nuances of how couples experiencing different types of violence react to couple therapy. Therefore, despite the benefits observed in the studies considered, we concluded with a recommendation to consider using couples therapy only after careful assessment for the suitability of couple therapy for that particular couple, and after taking necessary precautions to ensure the safety of both partners.

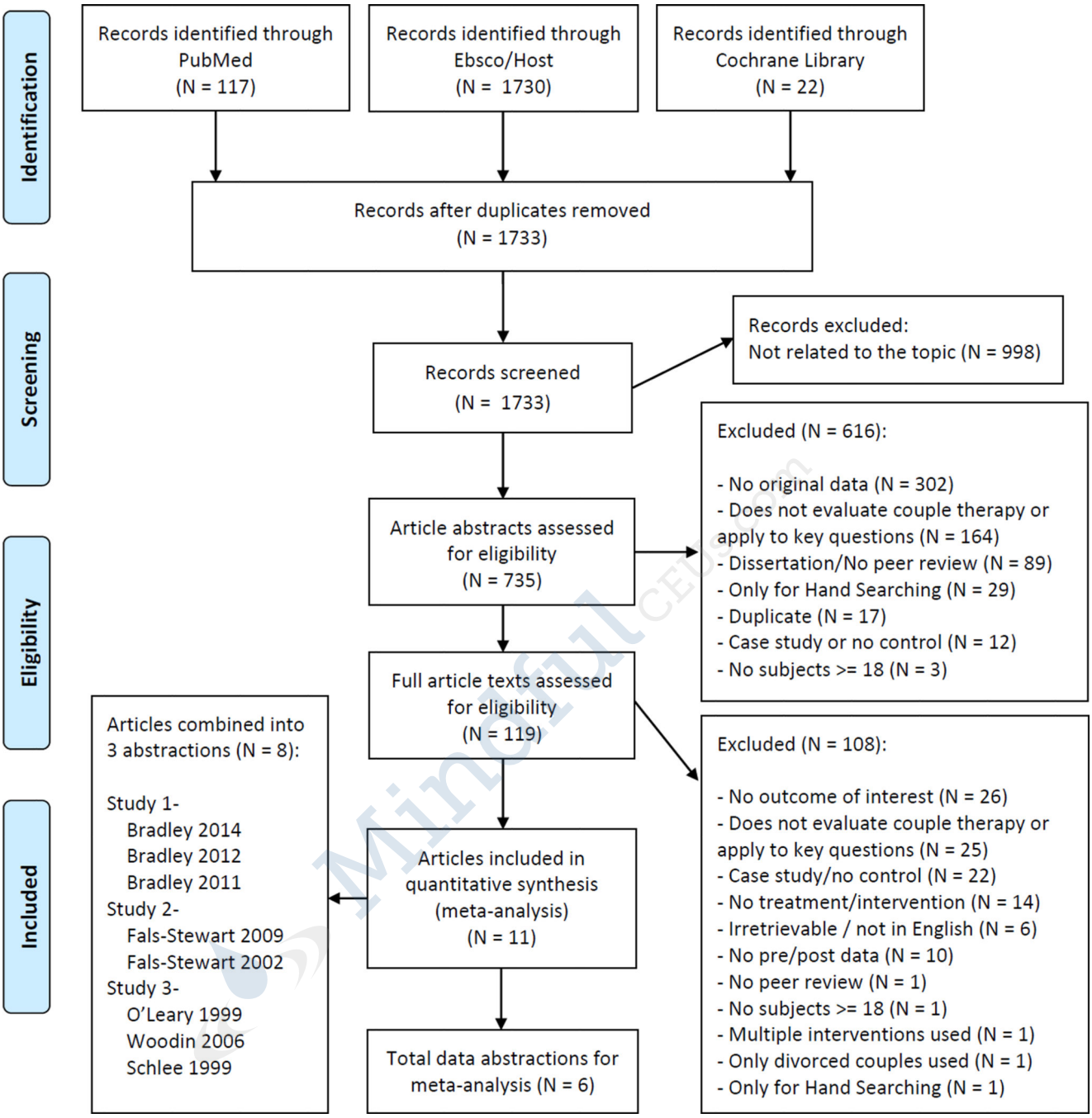

Figure 1.
PRISMA 2009 Flow Diagram for systematic review.

	Selection Bias		Performance Bias		Detection Bias		Attrition Bias		Reporting Bias	
	Q1	Q2	Q1	Q2	Q1	Q2	Q1	Q2	Q1	Q2
Bradley 2014 / Bradley 2012 / Bradley 2011	+	+	+	+	+	+	?	+	+	+
Fals-Stewart 2006	+	+	+	+	+	+	?	?	+	+
Fals-Stewart 2009 / Fals-Stewart 2002	+	+	+	+	+	+	?	?	+	+
Lam 2009	+	+	+	+	+	+	?	+	+	+
O'Leary 1999 / Woodin 2006 / Schlee 1998	+	+	+	+	+	+	-	+	+	+
Stith 2004	?	+	?	+	+	+	?	+	?	?

Figure 2.
Risk of bias for each study included in the analysis, evaluated by two reviewers, Q1 and Q2. Green (+) indicates low risk, yellow (?) indicates unclear risk, and red (–) indicates high risk.

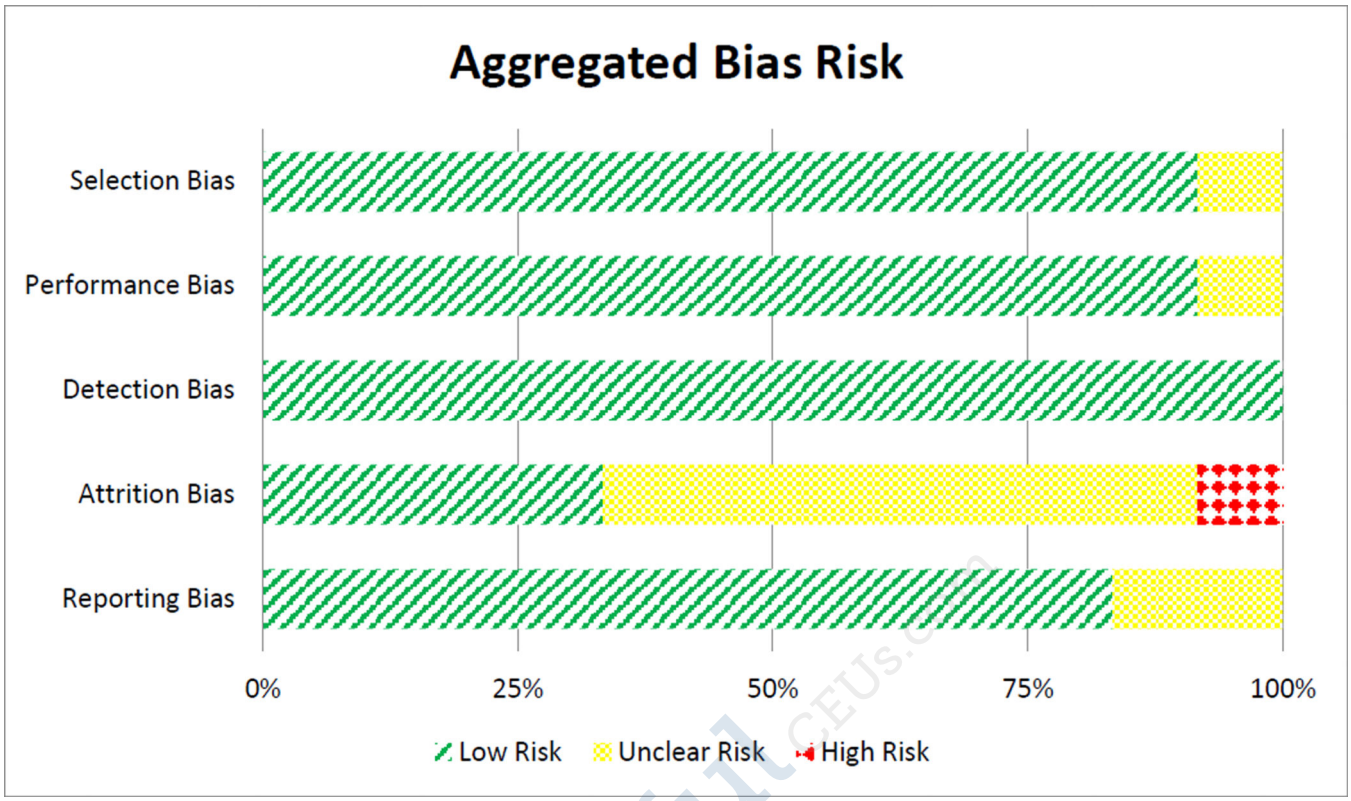

Figure 3.
Aggregated risk of bias overall for studies included in the analysis.

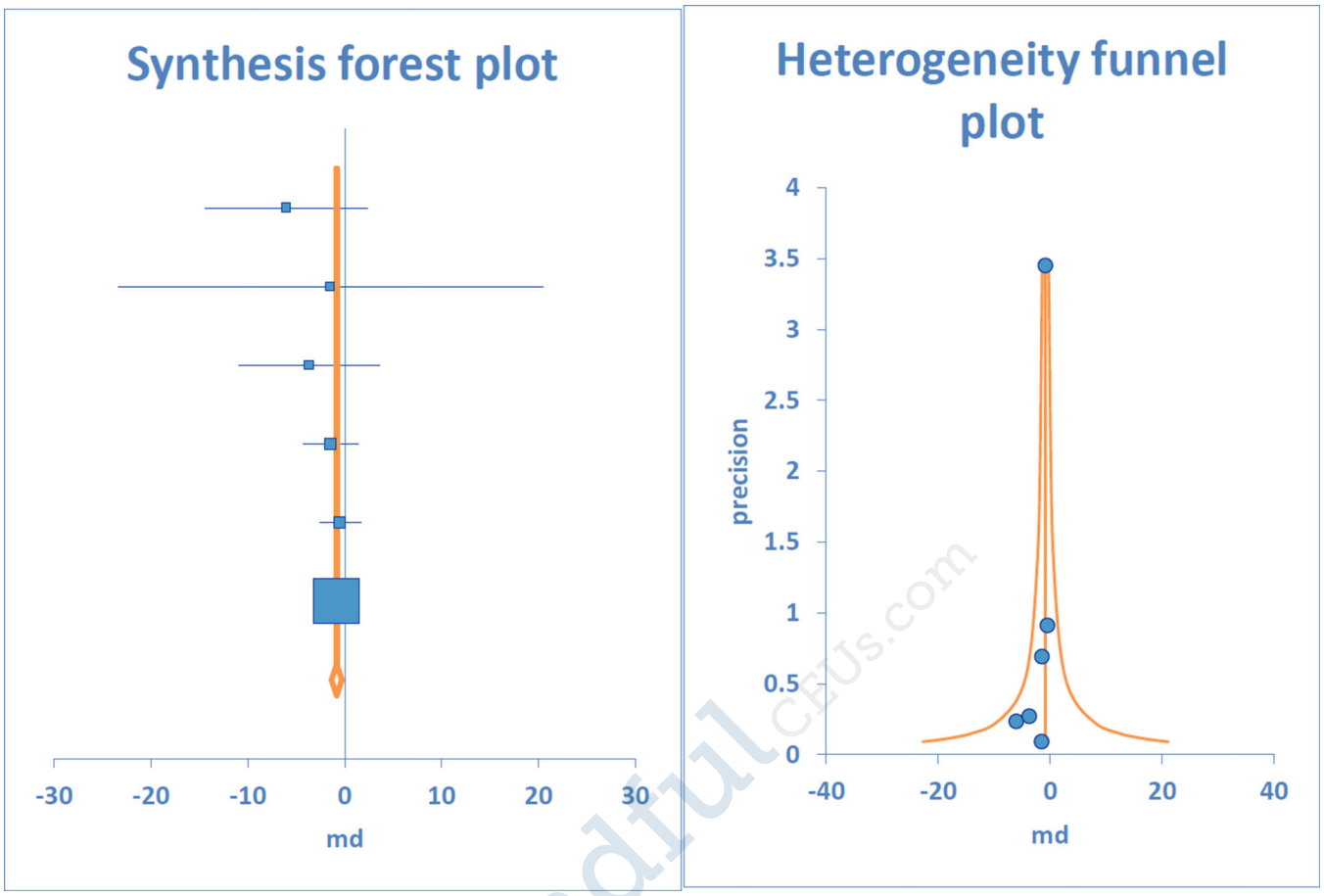

Figure 4.
Synthesis and Heterogeneity

Table 1

Characteristics of Included Studies

Study	Setting	Participants	Intervention	Control	Outcomes used in our analysis
Bradley 2014 Bradley 2012 Bradley 2011	Community Mental Health Agency	Couples over 18 in committed relationship experiencing IPV	Conjoint Group / Individual Couples therapy (CHRP)	No Treatment Control	Reductions in male and female perpetrated violence, measured by CTS2
Fals-Stewart 2006	Outpatient	Couples over 20 in a committed relationship with alcoholic female partner entering treatment	Individual Couples therapy (BCT)	Gender Specific Individual Therapy (IBT)	Reductions in male and female perpetrated violence, measured by TLFB-SV
Fals-Stewart 2009 Fals-Stewart 2002	Community Mental Health Agency	Couples over 20 in a committed relationship with substance abusing male partner entering treatment	Individual Couples therapy (BCT)	Gender Specific Group Therapy (IBT)	Reductions in male perpetrated violence (all and severe), measured by TLFB-SV
Lam 2009	Outpatient	Parents with male partner entering outpatient alcohol treatment	Individual Couples therapy (BCT)	Gender Specific Individual Therapy (IBT)	Reductions in male and female perpetrated violence, measured by TLFB-SV
O'Leary 1999 Woodin 2006 Schlee 1999	Not Reported	Couples in a committed relationship experiencing male perpetrated IPV	Conjoint Group Therapy (PACT)	Gender Specific Group Therapy (GST)	Reductions in male and female perpetrated violence (all, minor, and severe), measured by mCTS
Stith 2004	Not Reported	Couples over 18 in a committed relationship experiencing IPV	Individual Couples therapy (DVFCT)	No Treatment Control	Reductions in male and female perpetrated violence (all, minor, and severe) and total couple violence, measured by CTS2

BCT = Behavioral Couples Therapy
CHRP = Creating Healthy Relationships Program
DVFCT = Domestic Violence Focused Couples Treatment
IBT = Individually Based Therapy/Treatment
PACT = Physical Aggressive Couples Treatment
Revised Conflict Tactics Scales = CTS2
Modified Conflict Tactics Scales = mCTS
Timeline Followback Interview - Spousal Violence = TLFB-SV

Table 2

Meta-analysis Results

K	**N**	**wmd**	**ci−**	**ci+**	**z**	**p**	
6	**470**	**−0.84**	**−1.37**	**−0.30**	**−3.07**	**0.00**	
Id	**N**	**md**	**ci−**	**ci+**	**z**	**p**	**w**
Stith 2004	14	−6.02	−14.39	2.35	−1.41	0.16	0.41%
Bradley 2014, 2012,2011	100	−1.48	−23.39	20.43	−0.13	0.89	0.06%
O'Leary1999, Woodin2006, Schlee1998	37	−3.69	−10.97	3.59	−0.99	0.32	0.54%
FalsStewart2006	92	−1.44	−4.27	1.39	−1.00	0.32	3.58%
Lam2009	20	−0.45	−2.60	1.70	−0.41	0.68	6.21%
FalsStewart2009, 2002	207	−0.80	−1.37	−0.23	−2.76	0.01	89.20%

id = study identification tag during analysis
k = number of studies included in analysis
N = number of participants
wmd = weighted mean difference
md = value of male violence recidivism by point estimate
ci− = low end confidence interval value
ci+ = high end confidence interval value
z = z-score measuring relationship to mean
p = p-value for statistical significance
w = weight of study based on percentage of total combined participants

Mindful
Continuing Education

"This course was developed from the open access article: Giulia Ferrari, Roxane Agnew-Davies, Jayne Bailey, Louise Howard, Emma Howarth, Tim J. Peters, Lynnmarie Sardinha & Gene Solomon Feder (2016) Domestic violence and mental health: a cross-sectional survey of women seeking help from domestic violence support services, Global Health Action, 9:1, (DOI:10.3402/gha.v9.29890), used under the Creative Commons Attribution License."

"This course was developed from the open access article: Cognitive behavioural group therapy for male perpetrators of intimate partner violence: a systematic review, BMC Psychiatry 2019,19:11, (DOI:10.1186/s12888-019-2010-1), used under the Creative Commons Attribution License."

"This course was developed from the open access article: Couple Therapy for Intimate Partner Violence: A Systematic Review and Meta-Analysis, Karakurt et al., J Marital Fam Ther. 2016 October ; 42(4): 567–583. (DOI:10.1111/jmft.12178), used under the Creative Commons Attribution License."