Child Abuse & Neglect 36 (2012) 308–322

Contents lists available at SciVerse ScienceDirect

Child Abuse & Neglect

Integrated programs for mothers with substance abuse issues and their children: A systematic review of studies reporting on child outcomes☆

Alison Niccols[a,*], Karen Milligan[b], Ainsley Smith[c], Wendy Sword[d], Lehana Thabane[e], Joanna Henderson[f]

[a] 280 Holbrook Building, McMaster Children's Hospital-Chedoke Site, Box 2000, McMaster University, Hamilton, Ontario, Canada L8N 3Z5
[b] Integra, Toronto, Ontario, Canada M5P 1B9
[c] 284 Holbrook Building, McMaster Children's Hospital-Chedoke Site, Box 2000, McMaster University, Hamilton, Ontario, Canada L8N 3Z5
[d] School of Nursing HSC 3H48B, McMaster University, Hamilton, Ontario, Canada L8N 3Z5
[e] Centre for Evaluation of Medicines, McMaster University, Hamilton, Ontario, Canada L8N 1G6
[f] Child, Youth, and Family Program, Centre for Addiction and Mental Health, Toronto, Canada M5T 1R8

ARTICLE INFO

Article history:
Received 12 March 2011
Received in revised form
29 September 2011
Accepted 12 October 2011
Available online 5 April 2012

Keywords:
Children
Mothers
Substance abuse
Integrated programs

ABSTRACT

Background: Integrated treatment programs (those that include on-site pregnancy-, parenting-, or child-related services with addiction services) were developed to break the intergenerational cycle of addiction, potential child maltreatment, and poor outcomes for children.

Objectives: To examine the impact and effects of integrated programs for women with substance abuse issues and their children, we performed a systematic review of studies published from 1990 to 2011.

Methods: Literature search strategies included online bibliographic database searches, checking printed sources, and requests to researchers. Studies were included if all participants were mothers with substance abuse problems at baseline; the treatment program included at least 1 specific substance use treatment and at least 1 parenting or child treatment service; the study design was randomized, quasi-experimental, or cohort; and there were quantitative data on child outcomes. We summarized data on child development, growth, and emotional and behavioral outcomes.

Results: Thirteen studies (2 randomized trials, 3 quasi-experimental studies, 8 cohort studies; $N = 775$ children) were included in the review. Most studies using pre-post design indicated improvements in child development (with small to large effects, $ds = 0.007–1.132$) and emotional and behavioral functioning (with most available effect sizes being large, $ds = 0.652–1.132$). Comparison group studies revealed higher scores for infants of women in integrated programs than those not in treatment, with regard to development and most growth parameters (length, weight, and head circumference; with all available effect sizes being large, $ds = 1.16–2.48$). In studies comparing integrated to non-integrated programs, most improvements in emotional and behavioral functioning favored integrated programs and, where available, most effect sizes indicated that this advantage was small ($ds = 0.22–0.45$).

☆ The Canadian Institutes for Health Research (CIHR) provided funding for this project.

* Corresponding author.

doi:10.1016/j.chiabu.2011.10.007

Conclusions: Available evidence supports integrated programs, as findings suggest that they are associated with improvements in child development, growth, and emotional and behavioral functioning. More research is required comparing integrated to non-integrated programs. This review highlights the need for improved methodology, study quality, and reporting to improve our understanding of how best to meet the needs of children of women with substance abuse issues.

Introduction

Maternal substance abuse is a serious problem for the child welfare system. Estimates suggest that 50–80% of child welfare cases involve a parent who abuses alcohol or other drugs and mothers make up the majority of substance-abusing parents in the child welfare system (Substance Abuse and Mental Health Services Administration [SAMHSA], 2002; US Department of Health and Human Services, 1999). In the USA, 59% of the adults in substance abuse programs are parents (i.e., over 1 million of the 1.84 million in treatment) and 27% (294,000) have had 1 or more children removed by child welfare services (SAMHSA, 2002). Other estimates suggest that up to 70% of women in substance abuse treatment have children (US Department of Health and Human Services, 1999). These circumstances have serious implications for child welfare, health, and development and represent considerable human and financial burden to society.

Rates of substance abuse in women have been increasing (Greenfield, 2002) and substance abuse in women also is associated with a unique constellation of risk factors and needs, including greater vulnerability to adverse physiological consequences than men, greater prevalence of mental health problems, histories of physical or sexual abuse, serious medical problems, poor nutrition, relationship problems (including domestic violence), and deficits in social support (Hans, 1999; Hernandez-Avila, Rounsaville, & Kranzler, 2004). The unique risk factors and presenting needs of women have resulted in the development of women-specific comprehensive treatment models (Greenfield, 2002). However, in addition to having gender-specific needs, women with substance abuse issues also have unique needs as mothers.

Use of alcohol and other drugs during pregnancy can have profound effects on child health and development. Children born to women who used substances during pregnancy are at greater risk for prematurity, impaired physical growth and development, and physical and mental health problems (Barnard & McKeganey, 2004). The postnatal environment of children born to women with substance abuse issues also puts them at high risk for poor outcomes. Research has shown that women who abuse substances may have difficulties providing stable, nurturing environments for their children compounded by challenging life circumstances, including severe economic and social problems, such as lack of affordable housing and homelessness (Kelley, 1998). Maternal substance abuse has been associated with child neglect and abuse (Dunn et al., 2002) and the living conditions of children of women who abuse substances put them at high risk for family disruption, exposure to violence, and poor physical, academic, and socio-emotional outcomes (Conners et al., 2003). In 1 study of preschool age children of mothers with substance abuse issues, cognitive limitations were diagnosed in 69%, emotional or behavioral problems in 16%, and medical problems in 83% (Shulman, Shapira, & Hirshfield, 2000). At school-age, children of women who abuse substances have higher rates of emotional and behavioral problems than other children (Wilens, Biederman, Kiely, Bredin, & Spencer, 1995) and, in adolescence, are more likely to have a major psychiatric disorder (65%; Luthar, Cushing, Merikangas, & Rounsaville, 1998) and use alcohol and other drugs themselves (Legrand, Iacono, & McGue, 2005).

As maternal substance use is a growing problem for the child welfare system and society in general, there is an urgent need to identify effective interventions. Treatment for mothers with substance abuse issues and their children may represent an important opportunity for breaking the intergenerational cycle of addiction and dysfunction, improving child outcomes, and reducing costs (in terms of foster care placement, etc.). However, women with substance abuse issues report difficulties using conventional systems of care (for reasons including fear of losing custody of children, guilt, stigma, and lack of transportation), and request comprehensive services provided in a caring, "one-stop" setting (Howell & Chasnoff, 1999). Given the barriers, risks, and outcome implications, researchers, clinicians, and policy makers recommend that substance abuse treatment programs address women's needs as well as their children's needs through comprehensive, integrated services in centralized settings for both women and children (e.g., Howell & Chasnoff, 1999). This recognition has resulted in the development of numerous integrated treatment programs (those that include on-site pregnancy-, parenting-, or child-related services with addiction services), both residential and outpatient. Integrated residential programs or "therapeutic communities" offer long-term (15–18 months) treatment services to women and their children. Both types of programs typically are comprehensive and include group and individual addiction treatment, maternal mental health services, trauma treatment, parenting education and counseling, life skills training, prenatal education, medical and nutrition services, education and employment assistance, child care, children's services, and aftercare.

To date, no systematic reviews of studies of child outcomes of integrated programs have been conducted. Gender specific (i.e., women only) substance use treatment was examined in 1 systematic review and 1 meta-analysis. In their systematic review of 38 studies, Ashley, Marsden, and Brady (2003) found that programs with prenatal care or child care were associated with improved outcomes (substance use, mental health, birth outcomes, employment, and health). Similarly, in their meta-analysis, Orwin, Francisco, and Bernichon (2001) concluded that enhancing women-only addiction treatment programs with

prenatal care or therapeutic child care added value above and beyond the effects of standard women-only programs. Neither of these reviews specifically focused on integrated programs or examined child outcomes.

We examined the impact of integrated programs on child outcomes (child development, growth, and emotional and behavioral functioning) in a systematic review of relevant studies. The specific research questions guiding this systematic review were: (1) What is the impact of integrated programs on child outcomes from intake to post-test?; (2) Are integrated programs more effective than no treatment in improving outcomes for children?; and (3) Are integrated programs more effective than non-integrated programs in improving outcomes for children?

Method

Information sources and literature search

We used 3 main strategies to identify outcome studies of intervention programs for women with substance abuse issues and their children: online bibliographic database searches, checking printed sources, and requests to researchers (cf., Rosenthal, 1991). First, we searched relevant bibliographic databases (PsycINFO, MedLine, PubMed, Web of Science, EMBASE, Proquest Dissertations, Sociological Abstracts, and CINAHL) for studies published in English from 1990 to May 2011, using a subject heading and keyword search for the terms "substance abuse (or substance use or addict* or alcohol*) and intervention (or treatment or therapeutic or rehab*) and women (or mother) and child (or infant) and mental health and prenatal (or parent*), singly and in combination. Secondly, we examined reference lists of retrieved articles for potentially relevant documents. In addition, we manually searched relevant journals in the area (*Addiction, Addictive Behaviors*, *International Journal of the Addictions*, *Journal of Drug Issues*, *Journal of Psychoactive Drugs*, *Journal of Substance Abuse*, *Journal of Substance Abuse Treatment*, *Journal of Substance Use*, *and Substance Use and Misuse*). Documents that appeared to be relevant on the basis of titles or abstracts were retrieved. Finally, we searched for gray literature (technical reports, clinical trials registry, unpublished data). All researchers identified through these searches, as well as researchers presenting at relevant conferences identified using Google and Cross Currents (Upcoming Events), were contacted by to request any relevant published or unpublished data. Of the 200 researchers identified and emailed, 48% responded and 28 additional studies were identified.

Eligibility criteria and study selection

Eligibility criteria were based on our working definition of integrated programs being substance abuse treatment programs that provide comprehensive services that address substance abuse as well as maternal and child well-being through prenatal services, parenting programs, child care, or other child-centred services in a centralized setting. Our criteria were purposefully broad in order to include as many potentially relevant studies with outcome data as possible. Therefore, we included studies in our larger systematic review and meta-analysis if all of the following criteria were met:

(1) all participants were women who were pregnant or parenting,
(2) all participants had substance abuse problems at baseline. We included any study that reported that the participants had a diagnosis of chemical abuse and dependence. However, given that diagnostic or standardized measures of substance use were not routinely used at intake, enrollment in substance abuse treatment was considered to reflect substance abuse at a level that impacted daily functioning. Therefore, we included studies where all participants had specific substance abuse treatment for abuse of any drug (e.g., cocaine, crack, heroin, marijuana) or alcohol. We excluded any study that included only non-users or those at risk for substance use, and we contacted authors if the issue was unclear or not reported,
(3) the treatment program included at least 1 specific substance use treatment (e.g., individual or group therapy, methadone) and at least 1 parenting or child (<16 years) treatment service (e.g., prenatal care, child care, parenting classes),
(4) the substance use treatment was not only a smoking cessation program,
(5) the program was not for men or for women who were not pregnant or parenting,
(6) the study design was randomized, quasi-experimental, or cohort, and
(7) there were quantitative data on child outcomes or other outcomes (length of stay, treatment completion, maternal substance use, maternal well-being, or parenting).

Data extraction

Upon completion of the literature search, we developed a codebook based on theoretical treatment models, literature review, and data availability. We collected data on dependent variable characteristics (type of outcome measure, type of data), and outcome statistics (e.g., *F* value, *p* value) and coded study context (author, document date, type of document, country), methodology (sample size, attrition, study design), participant characteristics (age, marital status, education, employment, income, substance abuse history, previous substance abuse treatment, mental and physical health, involvement with the legal system), child characteristics (age, custody, involvement with child protection services, positive toxicology at birth), and treatment program characteristics (population served, planned length of treatment, intensity of treatment, location,

services). Project staff and investigators pilot tested the codebook, which we revised based on consensus before formally coding the studies. In a coding policy manual, we recorded variables that were added or deleted, decisions, and clarification of specific variables.

A trained research assistant (AS) coded each study and met frequently with the principal investigator (KM) during the development of the codebook and early stage of coding. Both AS and KM coded 20% of studies. We calculated Cohen's *Kappa* and percent agreement for all variables. There was 100% agreement for identification of dependent variables and, for client and program variables, 94% mean agreement for continuous variables and a *Kappa* of 0.97 for categorical variables. We resolved discrepancies by consensus.

There were considerable missing data (especially on client characteristics and program services) and limited quantitative data on outcomes (e.g., standard deviations, sample sizes). In an attempt to obtain missing data, we contacted 89 researchers up to 3 times each throughout the coding process. In total, 79% responded, with 37% providing some additional data.

Study quality

To assess the quality of randomized trials, we used the Jadad Scale (Moher, Jadad, & Tugwell, 1996), which is widely used in the medical literature. On the Jadad Scale, studies are rated on a scale from 0 to 5, with the highest possible score (5) given for those with descriptions of the randomization process, an appropriate method of randomization, double-blinding, an appropriate method of double-blinding, and withdrawal and dropouts. To assess the quality of non-randomized studies, we used the Newcastle-Ottawa Scale (NOS, Wells et al., n.d.). On the NOS, studies are rated on a scale from 0 to 9 on the basis of 3 main issues: study group selection, group comparability, and outcome ascertainment. NOS content validity and inter-rater reliability have been established and further evaluation is being conducted (Wells et al.). A trained research assistant (AS) and Master's student (JL) coded study quality. Inter-rater reliability, based on 16% (19) of the 120 eligible studies, was high, *Kappa* = 0.81. We resolved discrepancies by consensus.

Calculating effect sizes

In order to facilitate our summary and comparison of studies, effect sizes were calculated, where possible. We transformed results from each study to the standardized mean difference (Cohen's *d*) and used conventional definitions of effect size (*d*), for example, small = 0.20 or less (such as, one fifth or less of one standard deviation difference), medium = 0.50, large = 0.80 (such as, four fifths or more of one standard deviation difference) (Cohen, 1988).

Results

Study selection

In total, 328 studies were retrieved and coded for eligibility. Using the eligibility criteria, we excluded 63 studies because the study design was not randomized, quasi-experimental, or cohort and 50 because participants were not pregnant or parenting. In addition, 41 had no quantitative data on maternal or child outcomes, 27 had no substance use treatment, 16 had no pregnancy, parenting, or child service, 6 included men, and 5 included participants without a substance use problem at baseline. Therefore, of the 328 retrieved studies, we considered 120 studies eligible for inclusion in the larger systematic review and meta-analysis. Based on a random sample of 20% of the studies, inter-rater reliability for eligibility coding was high, *Kappa* = 0.81. We resolved discrepancies by consensus.

We estimated the completeness of the search using the capture re-capture method (Bennett, Latham, Stretton, & Anderson, 2004). Based on this method, the estimated number of missing articles is 8 (95% confidence interval [CI]: 2, 24), which suggests a 90% capture rate (i.e., the identified studies cover 90% of the search horizon). This reasonably high capture rate suggests that we retrieved a sufficient number of studies to avoid bias in the results of the systematic review. Of the 120 eligible studies, 107 studies did not have quantitative data on child outcomes. Therefore, for the present review, we included 13 studies. See Fig. 1 for a flow diagram.

Study characteristics

Studies varied in terms of study design, assessment times, and child outcome measures. In terms of study design, there were 8 cohort studies, one quasi-experimental study comparing an integrated program to no treatment, one quasi-experimental study comparing an integrated program to no treatment and nonusers, one quasi-experimental study comparing an integrated program to non-integrated child-specific treatment, and 2 randomized trials comparing integrated to non-integrated programs. Child outcomes were assessed at varying time points (e.g., at 3, 6, 12, 18, 24 months old, 3 or 6 months after intake, discharge, 6 or 12 months after discharge). Eight studies included measures of child development (Bayley Scales of Infant Development, Early Social Communication Scales, Denver Developmental Screening Test II, Child Development Inventory, Infant Neurological International Battery, Battelle Developmental Inventory), 2 studies examined growth parameters (length, weight, head circumference), and 5 studies included measures of child emotional and behavioral

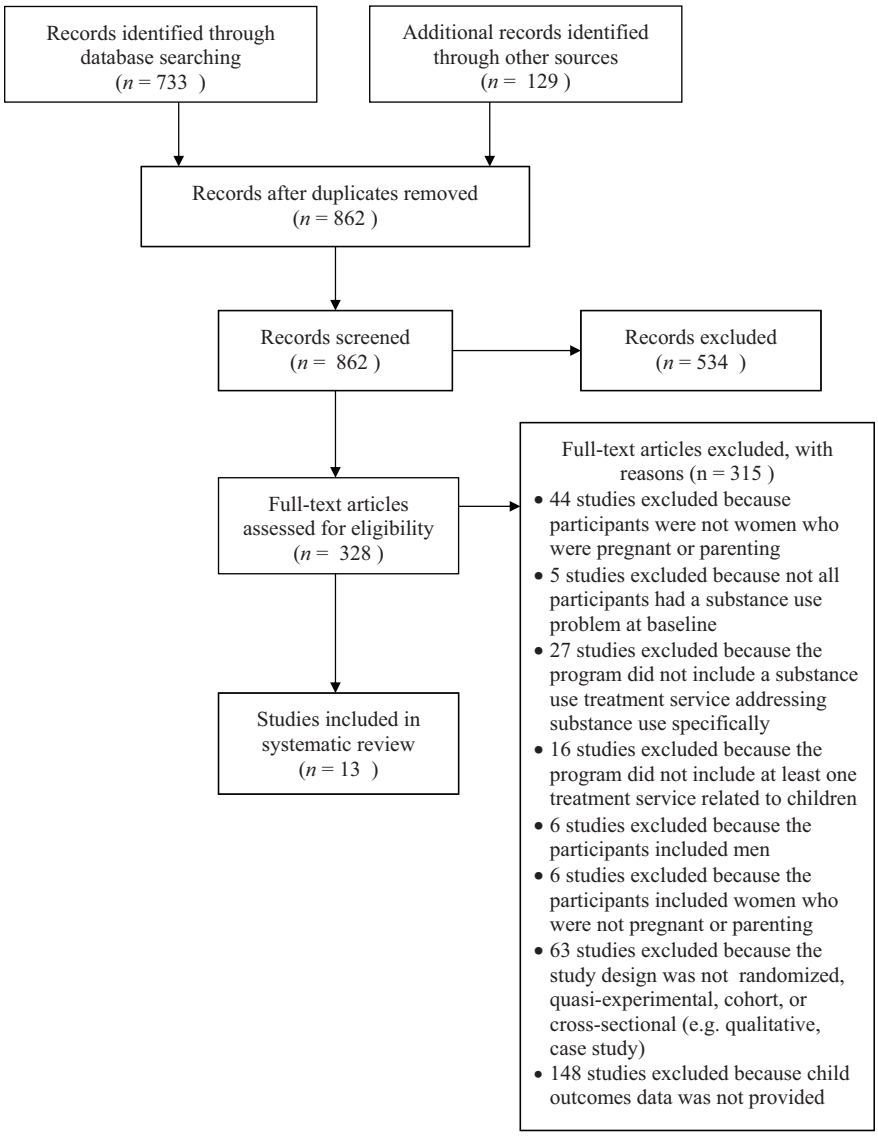

Fig. 1. Flow diagram of studies screened, assessed, and included.

functioning (Behavioral Assessment System for Children, Children's Depression Inventory, Behavioral and Emotional Rating Scale, Child Behavior Checklist).

Four of the 8 cohort studies were evaluations of integrated outpatient programs, 2 in the USA (Belcher et al., 2005; Jansson et al., 1996) and 2 in Canada (Niccols & Sword, 2005; Pepler, Moore, Motz, & Leslie, 2002). The other 4 studies were evaluations of integrated residential programs in the USA (Camp & Finkelstein, 1995; Conners, Bradley, Whiteside, Mansell, & Crone, 2001; Kerwin, 2007; Killeen & Brady, 2000). All programs were comprehensive and included group and individual addiction treatment, maternal mental health services, group and individual parenting education and counseling, life skills training, prenatal education, medical and nutrition services, education and employment assistance, obstetrical and pediatric care, child care, children's services, and aftercare. Women were pregnant or parenting and some women had other children (an average of 1 or 2). Mothers were typically poly-substance users whose primary substance problem was cocaine, methadone, heroin, alcohol, or cannabis. Their average age was 27–31 years and most had experienced trauma (such as childhood maltreatment), had mental health problems (such as clinical depression), and were unemployed, single mothers. Some programs had a large proportion of African-American clients and others had a large proportion of Caucasian clients, depending on the geographic setting. Four of the 8 cohort studies provided information on the percent of study participants who were involved with child protection services (47%— Camp & Finkelstein, 1995; 33%— Killeen & Brady, 2000; 69%— Niccols & Sword, 2005; 85%— Pepler et al., 2002). Three studies targeted infants 0–24 months, 3 were studies of preschoolers, and 7 were studies of older children. Programs were 6–12 months and had a high dropout rate (many > 50%).

In a quasi-experimental study with a total sample size of 126, Field et al. (1998) compared developmental outcomes and growth markers for infants of polydrug-using adolescent mothers participating in a comprehensive integrated outpatient program to those of infants of polydrug-using adolescent mothers not in treatment and non-using adolescent mothers (sample size was not specified by group). In another quasi-experimental study, Whiteside-Mansell, Crone, and Connors (1999) compared Bayley Scales of Infant Development Mental and Psychomotor Development Index scores and growth markers for children of women attending the Arkansas Center for Addictions Research, Education, and Services (AR-CARES) comprehensive residential integrated program to children of women who refused treatment. One-third of the participants were involved with child protection services. In the third quasi-experimental study, Noether et al. (2007) examined children (5–10 years old at intake) of women receiving services at 1 of 4 sites (2 residential and 2 outpatient) in the USA, comparing comprehensive integrated treatment services for mothers with substance abuse issues, mental health disorders, and histories of violence to child-specific treatment as usual (individual, group, and family services) as part of a larger study. One quarter of the women did not have custody of the target child. Integrated program services varied from site to site, however key characteristics were outreach, screening and assessment, substance abuse treatment, mental health services, trauma treatment, parenting skills training, and resource coordination.

In 1 of the 2 randomized trials, Luthar and Suchman (2000) randomly assigned mothers (of children under 16 years old) at 3 methadone clinics to standard treatment or standard treatment plus a mothers' group. Standard treatment included addiction counseling, pharmacological intervention (methadone), case management to assist with basic needs such as housing, welfare benefits, and legal aid. Psychosocial adjustment was assessed in children over 7 years old using the self-report and parent-report Behavioral Assessment System for Children. With another sample, Luthar, Suchman, and Altomare (2007) randomly assigned mothers (of children under 16 years old) at 3 methadone clinics to standard treatment plus recovery training or standard treatment plus a mothers' group. Children over the age of 7 were assessed for psychosocial adjustment using self-report and parent-report measures.

Study quality

Jadad Scale scores for the 2 randomized trials (Luthar & Suchman, 2000; Luthar et al., 2007) were both 3, which is a moderate score. Both studies were described as randomized but were not double blind as participants were aware of the treatment allocation. They both provided descriptions of an appropriate method of randomization and withdrawal and dropouts. NOS scores for the quasi-experimental studies (Field et al., 1998; Noether et al., 2007; Whiteside-Mansell et al., 1999) were 4, 3, and 2, respectively, which are low to moderate scores. NOS scores for the cohort studies (Belcher et al., 2005; Camp & Finkelstein, 1995; Conners et al., 2001; Jansson et al., 1996; Killeen & Brady, 2000; Niccols & Sword, 2005; Pepler et al., 2002) ranged from 1 to 3, which are low scores. It was unclear if low scores were due to poor study quality or poor reporting (e.g., if there was no description of the ascertainment of treatment exposure, then this item was scored as 0).

What is the impact of integrated programs on child outcomes from intake to post-test?

There were 6 cohort studies in which child development outcomes were reported. As can be seen in Table 1, all 9 mean developmental test scores for 6- to 24-month-old infants of women who participated in integrated programs were within or above one standard deviation of the normative mean in both studies in which they were examined. One cohort study reported that a large percentage (91–97%) of 6- and 12-month-old infants whose mothers participated in integrated programs scored over one standard deviation above the normative mean. In the 2 cohort studies reporting pre-post changes, most effects indicated positive child development outcomes for integrated programs, with effect sizes ranging from small to large ($ds = 0.007–1.132$). There were 2 cohort studies in which emotional and behavioral functioning outcomes were reported. As can be seen in Table 1, most effects indicated positive emotional and behavioral functioning outcomes for integrated programs from pre-test to post-test and, where available, most effect sizes were large ($ds = 0.652–1.132$). There were no studies examining the impact of integrated programs on child growth from intake to post-test.

Are integrated programs more effective than no treatment in improving outcomes for children?

There were 2 quasi-experimental studies in which child development outcomes and growth parameter outcomes (length, weight, and head circumference) were reported. As can be seen in Table 2, most developmental scores for 3-, 6-, and 12-month-old infants of women who participated in integrated programs were higher than those for infants of women not in treatment (and similar to those for infants of non-users). In the 1 study in which 18 month olds participated there were large differences between the groups, with children of women who participated in integrated programs scoring higher than children of women not in treatment. Also, most growth parameters for infants whose mothers participated in integrated programs were higher than those for infants of women not in treatment and, where available, all effect sizes were large ($ds = 1.16–2.48$). Below, we provide a narrative review of these 2 studies.

Field et al. (1998) compared developmental outcomes and growth markers for infants of polydrug-using adolescent mothers to those of infants of polydrug-using adolescent mothers not in treatment and non-using adolescent mothers (sample size not specified by group). At 3 and 6 months, scores on the Infant Neurological International Battery were similar to those for infants of non-users. At 12 months, children of women receiving treatment scored higher on the Early Social

Table 1
Studies of child outcomes (development and emotional and behavioral functioning) of integrated programs from intake to post-test.

Study	*n*	Measure	Time point	*ES* (*SE*)	*p*	*M* (*SD*)	% of sample	Quality
Belcher et al. (2005)	80	BSID MDI[a]	6 months old	[f]		100.9 (21.5)		1/9
			6–12 months old			103.2 (14.6)		
			12–18 months old			92.8 (6.5)		
Camp and Finkelstein (1995)	33	BSID MDI > 1	6 months old	[f]			97	1/9
	26	*SD* above *M*	12 months old				92	
	35	BSID PDI[b] > 1	6 months old	[f]			91	
	35	*SD* above *M*	12 months old				96	
Conners et al. (2001)	30	DDST II[c] Proportion with developmental delay	Follow up	0.449 (0.291)	0.123			2/9
Jansson et al. (1996)	24	BSID MDI	6 months old	[f]		100 (13.8)		3/9
	19		12 months old			107 (17.0)		
	2		24 months old			98 (6.8)		
	24	BSID PDI	6 months old	[f]		110 (13.3)		
	19		12 months old			107 (10.5)		
	2		24 months old			119 (11.5)		
Kerwin (2007)	11	CBCL Internalizing Behavior Problems	3 months after intake	0.487 (0.400)	0.223			n/a
			Discharge	0.346 (0.389)	0.374			
	11	CBCL Externalizing Behavior Problems	3 months after intake	0.785 (0.432)	0.069			
			Discharge	1.132 (0.484)	0.019			
	11	CBCL Total Behavior Problems	3 months after intake	0.652 (0.416)	0.117			
			Discharge	0.741 (0.427)	0.082			
Killeen and Brady (2000)		CBCL[d] Prosocial	Intake			60		2/9
	23	Behaviors Total Score	6-Month follow-up	[f]	<0.001	52		
		CBCL Prosocial	Intake			59.9		
	23	Behaviors Externalizing Score	6 month follow-up	[f]	<0.07	53.4		
		CBCL Prosocial	Intake			57.6		
	23	Behaviors Internalizing Score	6-Month follow-up	[f]	<0.001	50		
		CBCL	Intake			38.1		
	23	Behavioral Competence Subscale	6-Month follow-up	[f]	<0.05	45		

		CBCL School	Intake			39.9	
	23	Subscale	6-Month follow-up	[f]	ns	40.1	
		CBCL Social	Intake			38.1	
	23	Subscale	6-Month follow-up	[f]	<0.046	45.5	
Niccols and Sword (2005)	9	CDI[e] Social	3 months after intake	0.334 (0.343)	0.329		3/9
	7	Scale	6 months after intake	1.090 (0.477)	0.022		
	9	CDI Self-help	3 months after intake	0.406 (0.347)	0.241		
	7	Scale	6 months after intake	0.678 (0.419)	0.106		
	9	CDI Gross	3 months after intake	0.121 (0.335)	0.718		
	7	Motor Scale	6 months after intake	0.545 (0.405)	0.179		
	9	CDI Fine Motor	3 months after intake	−0.371 (0.345)	0.282		
	7	Scale	6 months after intake	0.007 (0.378)	0.986		
	9	CDI Expressive	3 months after intake	−0.086 (0.334)	0.796		
	7	Language Scale	6 months after intake	0.413 (0.394)	0.294		
	9	CDI Language	3 months after intake	−0.572 (0.360)	0.112		
	7	Comprehen-sion Scale	6 months after intake	0.220 (0.383)	0.564		
	9	CDI Letters	3 months after intake	0.104 (0.334)	0.755		
	7	Scale	6 months after intake	0.225 (0.383)	0.557		
	9	CDI Numbers	3 months after intake	0.135 (0.335)	0.686		
	7	Scale	6 months after intake	0.579 (0.408)	0.156		
Pepler et al. (2002)	25	Battelle Development Inventory	6 months after intake	0.847 (0.233)	0.000		2/9

[a] BSID MDI, Bayley Scales of Infant Development, Mental Development Index (normative $M = 100$, $SD = 15$).
[b] BSID PDI, Bayley Scales of Infant Development, Psychomotor Development Index (normative $M = 100$, $SD = 15$).
[c] DDST II, Denver Developmental Screening Test II.
[d] CBCL, Child Behavior Checklist (normative $M = 50$, $SD = 10$).
[e] CDI, Child Development Inventory.
[f] Unable to calculate effect size.

Table 2
Studies of child outcomes (development and growth) of integrated programs compared to no treatment.

Study	*n*	Groups	Design	Measure	Time point	*ES* (*SE*)	*p*	*M* diff[f] (*SD*)	Quality
Field et al. (1998)	Treatment: n/a Control: n/a *N* = 126	Outpatient integrated program vs. non-users	QE	ESCS[a] – Responding	12 months old	[e]		−0.2	4/9
				ESCS – Initiating	12 months old	[e]		−0.1	
				ESCS – Maintaining	12 months old	[e]		−0.2	
				BSID – MDI[b]	12 months old	[e]		−8.9	
				BSID – PDI[c]	12 months old	[e]		−9.0	
				INFANIB[d]	3 months old	[e]		−1.3	
					6 months old			−3.7	
				Length (in cm)	3 months old	[e]		−1.2	4/9
					6 months old			0.1	
					12 months old			−1.0	
				Weight (in g)	3 months old	[e]		−345.7	
					6 months old			−1224.9	
					12 months old			−730.9	
				Head circumference (in cm)	3 months old	[e]		−2.1	
					6 months old			−0.2	
					12 months old			−0.4	
Field et al. (1998)	Treatment: n/a Control: n/a *N* = 126	Outpatient integrated program vs. no treatment	QE	ESCS – Responding	12 months old	[e]		0.5	4/9
				ESCS – Initiating	12 months old	[e]		0.4	
				ESCS – Maintaining	12 months old	[e]		0.3	
				BSID – MDI	12 months old	[e]		10.3	
				BSID – PDI	12 months old	[e]		0.9	
				INFANIB	3 months old	[e]		−0.6	
					6 months old			1.5	
				Length (in cm)	3 months old	[e]		−0.7	
					6 months old			2.0	
					12 months old			0.9	
				Weight (in g)	3 months old	[e]		−103.4	
					6 months old			220.7	
					12 months old			302.1	
				Head circumference (in cm)	3 months old	[e]		−0.9	
					6 months old			2.5	
					12 months old			3.5	

Whiteside-Mansell et al. (1999)	Treatment: 16 Control: 3 Treatment: 9 Control: 5 Treatment: 6 Control: 1	Integrated program vs. no treatment	QE	BSID – MDI	6 months old 12 months old 18 months old	−0.17 (0.63) −0.43 (0.56)	0.77 0.45	12.7 (6, 0)	2/9
				BSID – PDI	6 months old 12 months old 18 months old	−0.37 (0.63) 0.96 (0.59)	0.56 0.10	18.5 (6.3, 0)	
	Treatment: 7 Control: 2 Treatment: 5 Control: 2 Treatment: 2 Control: 1			Length percentile	6 months old 12 months old 18 months old	1.16 (0.85)	0.17	45 (12.5) 27.5 (6.3, 0)	2/9
				Weight percentile	6 months old 12 month old 18 months old	1.16 (0.85) 2.48 (0.97)	0.17 0.01	15.0	
				Head circumference percentile	6 months old 12 months old 18 months old	1.82 (0.91) 2.36 (1.05)	0.05 0.02	32.5 (6.3)	

[a] ESCS, Early Social Communication Scale (normative *SD* = 1–2).
[b] BSID – MDI, Bayley Scales of Infant Development, Mental Development Index (normative *SD* = 15).
[c] BSID – PDI, Bayley Scales of Infant Development, Psychomotor Development Index (normative *SD* = 15).
[d] INFANIB, Infant Neurological International Battery (normative *SD* = 3.6).
[e] Unable to calculate effect size.
[f] Mean difference (treatment group mean – control group mean).

Communication Scales than children of women not in treatment, with scores approximating those of children of non-users. Children of women receiving treatment also scored higher on the Bayley Scales of Infant Development Mental Development Index than children of women not in treatment, but lower than the children of non-users. Scores on the Bayley Scales of Infant Development Psychomotor Development Index did not differ between children of women receiving or not receiving treatment and both of these groups scored lower than children of non-users. Mothers who participated in the integrated program had infants who were similar to non-users in growth parameters assessed when they were 3, 6, and 12 months old. Compared to no treatment, all growth parameter comparisons favored integrated treatment at 6 and 12 months, but not at 3 months. Unfortunately, statistical tests were not reported in this study.

Whiteside-Mansell et al. (1999) compared Bayley Scales of Infant Development Mental and Psychomotor Development Index scores and growth markers for children of women attending an integrated program to children of women who refused treatment. At 6 and 12 months, there were no significant differences between the groups in developmental scores, however, at 12 months there was a large effect size (0.96) for the Psychomotor Index favoring children whose mothers were in treatment. At 18 months, the treatment group outperformed the no treatment group in developmental scores, but small sample size precluded statistical comparison. At each age, all children of participating women scored in the normal range on the developmental tests whereas at 12 months 1 child of a non-participating woman scored below the normal range on both tests. At 6, 12, and 18 months of age, both groups of children were in the normal range for weight, length, and head circumference, with higher parameters for children in the integrated program than those in the no treatment comparison group.

There were no studies of child emotional and behavioral functioning outcomes comparing integrated programs to no treatment.

Are integrated programs more effective than non-integrated programs in improving outcomes for children?

There was one quasi-experimental study and 2 randomized trials in which emotional and behavioral functioning outcomes were reported. As can be seen in Table 3, most effects favored integrated over non-integrated treatment and, where available, most effect sizes indicated that this advantage was small (*d*s = 0.22–0.45). Below, we provide a narrative review of each of these 3 studies.

Luthar and Suchman (2000) randomly assigned mothers at 3 methadone clinics to standard treatment or standard treatment plus a mothers' group. Psychosocial adjustment was assessed in children over 7 years old using the self-report and parent-report Behavioral Assessment System for Children. At the end of the 6-month treatment and at 6-month follow-up, most (7/8) effects favored the integrated treatment, with small to medium effect sizes (*d*s = 0.22–0.53) that were not statistically significant (likely due to small sample size).

With another sample, Luthar et al. (2007) randomly assigned mothers at 3 methadone clinics to standard treatment plus recovery training or standard treatment plus a mothers' group. Children over the age of 7 were assessed for psychosocial adjustment using self-report and parent-report measures. There were no significant group differences on the parent-reported Behavioral Assessment System for Children. At the end of the 6-month treatment, scores for children whose mothers participated in the integrated program were significantly lower than the comparison group on the self-reported Behavioral Assessment System for Children and the Children's Depression Inventory. However, at 6-month follow-up, the circumstances reversed, with scores for children whose mothers participated in integrated treatment showing more maladjustment by self report than the comparison group. The authors highlight the need for continuity of care past formal completion of a treatment program.

Noether et al. (2007) examined children of women receiving services, comparing comprehensive integrated treatment services for mothers with substance abuse issues to child-specific treatment as usual. At both 6- and 12-month follow-up, scores on the Behavioral and Emotional Rating Scales showed more improvement for the children who participated in the integrated program than those of children in the control group. In addition, the authors found that significant predictors of positive child outcomes included younger child age, participation in a residential (vs. outpatient) program, having witnessed household violence, and non-Black race.

There were no studies of child development or growth outcomes comparing integrated to non-integrated programs.

Discussion

The purpose of this systematic review was to examine the effects of integrated treatment programs on child outcomes, guided by 3 research questions which we address below.

What is the impact of integrated programs on child outcomes from intake to post-test?

In the studies reviewed, developmental test scores for 6- to 24-month-old infants of women who participated in integrated programs were within or above one standard deviation of the normative mean. Most pre-post effects indicated small to large improvements in child development scores. With regard to emotional and behavioral functioning, most effects indicated

Table 3
Studies of child outcomes (emotional and behavioral functioning) of integrated programs compared to non-integrated programs.

Study	*n*	Groups	Design	Measure	Time point	ES (SE)	*p*	*M* diff[e] (*SD*)	Quality
Luthar and Suchman (2000)	Treatment: 32 Control: 20	Standard methadone treatment plus maternal psychotherapy vs. standard methadone treatment	RCT	BASC[a] – Child Maladjustment Scale – mother's report	Discharge 6-Month follow-up	0.29 0.45	>0.05 >0.05		3/5
				BASC – Clinical Maladjustment Scale – self report	Discharge 6-Month follow-up	0.22 0.39	>0.05 >0.05		
				BASC – Personal Adjustment Scale – self report	Discharge 6-Month follow-up	0.53 0.25	>0.05 >0.05		
				BASC – School Maladjustment Scale – self report	Discharge 6-Month follow-up	0.22 0.00	>0.05 >0.05		
Luthar et al. (2007)	Treatment: 60 Control: 67 Treatment: 50 Control: 58	Standard methadone treatment plus maternal psychotherapy vs. standard methadone treatment plus recovery training	RCT	BASC – Child Maladjustment Scale – mother's report	Discharge 6-Month follow-up	[d]	0.99 0.72	1.9 0.1	3/5
				BASC – Child Maladjustment Scale – self report	Discharge 6-Month follow-up	[d]	0.02 0.02	−2.3 3.5	
				Children's Depression Inventory[b]	Discharge 6-Month follow-up	[d]	0.04 0.12	−0.64 2.09	
Noether et al. (2007)	Treatment: 94 Control: 107 Treatment: 92 Control: 118	Integrated program vs. child treatment	QE	BERS[c] Strength Quotient	6-Month follow-up 12-Month follow up	[d]		1.0 2.8	3/9
				BERS – Relationship Tools Subscale	6-Month follow-up 12-Month follow up	0.318 0.431	0.028 0.007		
				BERS – Family Involvement Subscale	6-Month follow up 12-Month follow up	0.266 0.520	0.008 0.000		
				BERS – Capacity for Closeness Subscale	6-Month follow-up 12-Month follow up	n/a 0.328	>0.05 0.017		
				BERS – Positive Self Identify Subscale	6-Month follow-up 12-Month follow up	0.329 0.461	0.024 0.003		

[a] BASC, Behavioral Assessment System for Children (≥60 is considered clinical).
[b] Scores ≥ 12 are considered clinical.
[c] BERS, Behavioral and Emotional Rating Scale (normative *SD* = 15).
[d] Unable to calculate effect size.
[e] Mean difference (treatment group mean – control group mean).

improvements from pre-test to post-test and, where available, most effect sizes were large. Therefore, integrated programs appear to have a positive impact on child development and emotional and behavioral functioning from intake to post-test.

Are integrated programs more effective than no treatment in improving outcomes for children?

Most developmental scores for infants in integrated programs appeared higher than those for infants of women not in treatment and similar to those for infants of non-users. In terms of growth parameters (length, weight, and head circumference), most measures for infants in integrated programs were higher than those for infants of women not in treatment and, where available, all effect sizes were large. Therefore, integrated programs may be more effective than no treatment in improving child development and child growth.

Are integrated programs more effective than non-integrated programs in improving outcomes for children?

In comparison group studies of emotional and behavioral functioning, most effects favored integrated over non-integrated treatment and, where available, most effect sizes were small. As such, available evidence suggests that integrated programs may be associated with a small advantage over non-integrated programs in some child emotional and behavioral functioning outcomes. Unfortunately, there were no studies comparing integrated to non-integrated programs on child development and growth outcomes.

Implications

These findings are consistent with those reported in a meta-analysis of substance abuse treatment for women (Orwin et al., 2001), our own findings from a meta-analysis of integrated programs showing their positive impact on birth outcomes (Milligan et al., 2011), qualitative studies in which women reported that integrated programs improved their children's behavior and development (Sword et al., 2009), and the results of research on parent interventions with other at-risk populations (Cedar & Levant, 1990). Results from this systematic review are important given the risks for child maltreatment and poor outcomes in children of women with substance abuse issues (Dunn et al., 2002). The findings suggest that the risks to these children could be minimized with intervention, which could have long-term impact. For example, integrated programs may reduce emotional and behavioral problems in children, which has important implications for prevention of future psychiatric problems and substance abuse in these children (Killeen & Brady, 2000). Even though the advantage of integrated programs over non-integrated programs may be small, it could have a potentially large impact on the associated financial and human burden in this vulnerable population (e.g., it may reduce the need for foster care placement, child treatment, psychiatric admissions, crime).

Limitations

There were a number of challenges encountered in completing this systematic review that highlight current limitations in research on integrated treatment programs. First, among the 120 studies examining outcomes of integrated programs, there were only 13 with data on outcomes for children, despite the fact that improving child outcomes is often a stated goal of integrated programs. Among the studies reporting child outcome data, few were comparison group studies. Cohort studies provide information about child outcomes for those in integrated treatment but do not provide a comparison group that enables one to determine relative program effectiveness, for example, if these outcomes are significantly better than those for children of women who do not participate in treatment or who participated in other types of treatment. More comparison group studies are needed prior to concluding that there is robust support for the provision of these more comprehensive and population-specific services. Despite the limited number of studies included in the systematic review, we are confident that the search was not biased. We used several approaches to mitigate potential bias, including our attempts to identify gray literature by searching databases that include unpublished studies and contacting researchers for unpublished data as well as our use of the capture re-capture method to estimate the completeness of the literature search (identified studies covered 90% of the search horizon, suggesting that a sufficient number of studies were retrieved to avoid bias).

A second limitation of the present systematic review is that study quality was not high, as is typical of the substance abuse treatment field generally (Prendergast, Podus, Chang & Urada, 2002). Studies included in the systematic review were of low to moderate quality, although it was unclear if the scores reflected study design per se or the reporting of study quality elements. Although including lower quality studies weakens the strength of causal inferences that can be drawn from the synthesis, it increases the generalizability of the findings (Wilson, 2000).

A third limitation is that studies had small samples and relatively few child outcome measures. Few studies involved teacher-report measures of child functioning, which may be more objective, valid, and sensitive to change than self- or parent-report measures. Also, studies did not involve measures of some important areas of functioning that can be impacted by maternal substance abuse, such as infant attachment security, language abilities, attention, and health, nor did they involve longitudinal follow-up on growth and development and psychiatric, social, and academic functioning, or an assessment of cost effectiveness.

Fourth, missing data limited our review. Information on research methods and data needed to calculate effect sizes precluded meta-analysis and hampered attempts at assessing study quality. Often program, client, and study characteristics that might moderate treatment outcomes were not available. For example, only 6 of the 13 studies provided descriptive information on child protection services involvement of the study participants. Moderator analyses can have important practice implications, however, specific recommendations (e.g., regarding specific intervention strategies or specific subpopulations to target such as those in or out of foster care or younger or older children) await further research with better reporting to allow meta-analysis of variables that impact outcomes (cf., Prendergast et al., 2002).

Recommendations for future research

More high quality studies comparing non-integrated to integrated programs are needed, especially studies examining a variety of child outcomes. A multi-site study could address statistical limitations inherent in studies with small samples. The most rigorous design would be a randomized trial, but this may be challenging in the world of real-life service provision (e.g., program administrators likely would not agree to participate). Further, the examination of moderators is critical, given the variability in clients served and services offered in integrated and non-integrated substance abuse treatment programs. Also, examination of moderators would help identify effective components of intervention and ultimately plan studies to examine what works best for whom and under what circumstances. As some effects may not be immediately evident, follow-up for at least 2 years (or, ideally, longer) would be advisable. Given the challenges of following children longitudinally, particularly those involved with child protection services, collaboration with the child welfare system would be important to ensure representative longitudinal data. Linear regression, generalized estimating equations, or linear growth curve modelling could be used to analyze child outcomes with group and several other variables as predictors (e.g., demographic variables, maternal and child characteristics, program components), as well as the impact of mediators and moderators over time (Curran & Muthen, 1999). The propensity score method could be used to address the potential problem of baseline differences between groups (Rosenbaum & Rubin, 1983). Ensuring the availability of essential information to assess study quality and describe studies in future systematic reviews or meta-analyses could be accomplished by improvements in the editorial review process and creation of a registry of funded studies that would require submission of standard information (such as the Cochrane Collaboration on health care intervention), as has been recommended previously by researchers conducting substance abuse treatment meta-analyses (e.g., Prendergast et al., 2002).

Conclusion

The findings from this systematic review suggest that integrated programs for women with substance abuse issues and their children are associated with positive impacts on child development, growth, and emotional and behavioral functioning. These findings are encouraging in terms of the preventive potential for breaking the cycle of addiction, dysfunctional parenting, and poor outcomes for many vulnerable children. Consistent with the recommendations for research synthesis of Cooper, Hedges, and Valentine (2009), this review addresses an important and under-recognized, yet growing, area of research. The findings suggest the potential promise of integrated programs and highlight research gaps in study design, quality, and reporting practices. Future research involving prospective longitudinal studies with comparison group designs, larger samples, and full descriptions of the target population and the intervention program is recommended. To our knowledge, this systematic review is the first to examine the impact of integrated programs on child outcomes. Given that approximately one third of substance abusers are women of child-bearing age (World Health Organization, 2008), substance use among pregnant and parenting women is a serious problem for the child welfare system, and the burden of suffering due to maternal substance abuse is great, the findings from this review are noteworthy and support the need for more high quality research on integrated treatment programs for women with substance abuse issues and their children. The effectiveness of integrated programs warrant further exploration and investigation, as the implications of their wide-spread implementation may include reduced costs, increased access, and more positive outcomes for children.

Acknowledgment

The authors are grateful to research assistants, Jacky Chan, Joyce Li, and Jennifer Liu.

References

Ashley, O. S., Marsden, M. E., & Brady, T. M. (2003). Effectiveness of substance abuse treatment for women: A review. *American Journal of Drug and Alcohol Abuse*, *29*, 19–53.

Barnard, M. A., & McKeganey, N. P. (2004). The impact of parental problem drug use on children: What is the problem and what is being done to help? *Addiction*, *99*, 552–559.

Belcher, H., Butz, A., Wallace, P., Hoon, A., Reinhardt, E., Reeves, S., & Pulsifier, M. (2005). Spectrum of early intervention services for children with intrauterine drug exposure. *Infants and Young Children*, *18*, 2–15.

Bennett, C., Latham, N., Stretton, C., & Anderson, C. (2004). Capture–recapture is a useful method for assessing publication bias. *Journal of Clinical Epidemiology*, *57*, 349–357.

Camp, J. M., & Finkelstein, N. (1995). *Fostering effective parenting skills and healthy child development within residential substance abuse treatment settings.* Cambridge, MA: Center for Substance Abuse Prevention.

Cedar, B., & Levant, R. F. (1990). A meta-analysis of the effects of parent effectiveness training. *The American Journal of Family Therapy*, *18*, 373–384.
Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
Conners, N. A., Bradley, R. H., Mansell, L. W., Liu, J. Y., Roberts, T. J., Burgdorf, K., & Herrell, J. M. (2003). Children of mothers with serious substance abuse problems: An accumulation of risks. *American Journal of Drug & Alcohol Abuse*, *29*(4), 743–758.
Conners, N. A., Bradley, R. H., Whiteside-Mansell, L., & Crone, C. C. (2001). A comprehensive substance abuse treatment program for women and their children: An initial evaluation. *Journal of Substance Abuse Treatment*, *21*, 67–75.
Cooper, H., Hedges, L., & Valentine, J. (Eds.). (2009). *The handbook of research synthesis and meta-analysis*. New York, NY: Russell Sage.
Curran, P. J., & Muthen, B. O. (1999). The application of latent curve analysis to testing developmental theories in intervention research. *American Journal of Community Psychology*, *27*, 567–595.
Dunn, M. G., Tarter, R. E., Mezzich, A. C., Vanyukov, M., Kirisci, L., & Kirillova, G. (2002). Origins and consequences of child neglect in substance abuse families. *Clinical Psychology Review*, *22*, 1063–1090.
Field., T. M., Scafidi, F., Pickens, J., Prodromidis, M., Pelaez-Nogueras, M., Torquati, J., Wilcox, H., Malphurs, J., Schanberg, S., & Kuhn, C. (1998). Polydrug-using adolescent mothers and their infants receiving early intervention. *Adolescence*, *33*, 117–143.
Greenfield, S. (2002). Women & alcohol use disorders. *Harvard Review of Psychiatry*, *10*, 76–85.
Hans, S. L. (1999). Demographic and psychosocial characteristics of substance-abusing pregnant women. *Clinics in Perinatology*, *26*, 55–74.
Hernandez-Avila, C., Rounsaville, B., & Kranzler, H. (2004). Opioid-, cannabis- and alcohol-dependent women show more rapid progression to substance abuse treatment. *Drug and Alcohol Dependence*, *74*, 265–272.
Howell, E. M., & Chasnoff, I. J. (1999). Perinatal substance abuse treatment: Findings from focus groups with clients and providers. *Journal of Substance Abuse Treatment*, *17*, 139–148.
Jansson, L. M., Svikis, D., Lee, J., Paluzzi, P., Rutigliano, P., & Hackerman, F. (1996). Pregnancy and addiction: A comprehensive care model. *Journal of Substance Abuse Treatment*, *13*, 321–329.
Kelley, S. J. (1998). Stress and coping behaviors of substance-abusing mothers. *Journal of Social and Pediatric Nurses*, *3*, 103–110.
Kerwin, M. L. (2007). *Treatment outcome data for women in substance abuse treatment group*. Unpublished raw data.
Killeen, T., & Brady, K. T. (2000). Parental stress and child behavioral outcomes following substance abuse residential treatment: Follow-up at 6 and 12 months. *Journal of Substance Abuse Treatment*, *19*, 23–29.
Legrand, L. N., Iacono, W. G., & McGue, M. (2005). Predicting addiction. *American Scientist*, *93*, 140–147.
Luthar, S. S., Cushing, G., Merikangas, K., & Rounsaville, B. J. (1998). Multiple jeopardy: Risk/protective factors among addicted mothers' offspring. *Development and Psychopathology*, *10*, 117–136.
Luthar, S. S., & Suchman, N. E. (2000). Relational psychotherapy mothers' group: A developmentally informed intervention for at-risk mothers. *Development and Psychopathology*, *12*, 235–253.
Luthar, S., Suchman, E. S., & Altomare, M. (2007). Relational psychotherapy mother's group: A randomized clinical trial for substance abusing mothers. *Development and Psychopathology*, *19*, 243–261.
Milligan, K., Niccols, A., Sword, W., Thabane, L., Henderson, J., & Smith, A. (2011). Birth outcomes for infants born to women participating in integrated substance abuse treatment programs: A meta-analytic review. *Addiction Theory & Research*, *19*(6), 542–555.
Moher, D., Jadad, A. R., & Tugwell, P. (1996). Assessing the quality of randomized controlled trials. *International Journal of Technological Assessment of Health Care*, *12*, 195–208.
Niccols, A., & Sword, W. (2005). New choices for substance-using mothers and their children: Preliminary evaluation. *Journal of Substance Use*, *10*, 239–251.
Noether, C. D., Brown, V., Finkelstein, N., Russell, L. A., VanDeMark, N. R., Morris, L. A., & Graeber, C. (2007). Promoting resiliency in children of mothers with co-occurring disorders and histories of trauma: Impact of a skills based intervention program on child outcomes. *Journal of Community Psychology*, *35*, 823–843.
Orwin, R., Francisco, L., & Bernichon, T. (2001). *Effectiveness of women's substance abuse treatment programs: A meta-analysis*. Fairfax, VA: Center for Substance Abuse Treatment.
Pepler, D., Moore, T. E., Motz, M., & Leslie, M. (2002). *Breaking the cycle: The evaluation report (1995–2000)*. Toronto: Health Canada.
Prendergast, M. L., Podus, D., Chang, E., & Urada, D. (2002). The effectiveness of drug abuse treatment: A meta-analysis. *Drug and Alcohol Dependence*, *67*, 53–72.
Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of propensity score in observational studies for causal effects. *Biometrika*, *70*, 41–55.
Rosenthal, R. (1991). *Meta-analytic procedures for social research*. Newbury Park, CA: Sage.
Shulman, L. H., Shapira, S. R., & Hirshfield, S. (2000). Outreach developmental services to children of patients in treatment for substance abuse. *American Journal of Public Health*, *90*, 1930–1933.
Substance Abuse and Mental Health Services Administration. (2002). *Results from the 2001 National Household Survey on Drug Abuse*. NHSDA Series H-17, DHHS Publication No. SMA 02-3758. Rockville, MD: Office of Applied Studies.
Sword, W., Jack, S., Niccols, A., Milligan, K., Henderson, J., & Thabane, L. (2009). Integrated programs for women with substance use issues and their children: A qualitative meta-synthesis of processes and outcomes. *Harm Reduction Journal*, *6*, 32–37.
US Department of Health and Human Services. (1999). *1998 annual report to congress on the evaluation of the comprehensive community mental health services for children and their families program* (SAMHSA/CMHS unpublished draft manuscript). Rockville, MD: HHS Substance Abuse and Mental Health Services Administration.
Wells, G. A., Shea, B., O'Connell, D., Peterson, J., Welch, V., Losos, M., Tugwell, P. (n.d.). *The Newcastle-Ottawa Scale for assessing the quality of nonrandomized studies*. Retrieved from http://www.ohri.ca/programs/clinical_epidemiology/oxford.htm
Whiteside-Mansell, L., Crone, C. C., & Conners, N. A. (1999). The development and evaluation of an alcohol and drug prevention and treatment program for women and children. *Journal of Substance Abuse Treatment*, *16*, 265–275.
Wilens, T. E., Biederman, J., Kiely, K., Bredin, E., & Spencer, T. J. (1995). Behavioral and emotional disturbances in the high-risk children of parents with opioid dependence. *Journal of the American Academy of Child and Adolescent Psychiatry*, *34*, 779–785.
Wilson, D. B. (2000). Meta-analyses in alcohol and other drug abuse treatment research. *Addiction*, *95*, 419–438.
World Health Organization. (2008). *Principles of drug dependence treatment*. Geneva, Switzerland: World Health Organization.