Published in final edited form as:

Arch Womens Ment Health. 2016 April ; 19(2): 415–421. doi:10.1007/s00737-015-0536-x.

Test-retest reliability of retrospective self-reported maternal exposure to childhood abuse and neglect

Alison L. Cammack¹, Carol J. Hogue¹, Carolyn D. Drews-Botsch¹, Michael R. Kramer¹, Bradley D. Pearce¹, Bettina T. Knight², Zachary N. Stowe^{2,3,4}, and D. Jeffrey Newport⁵

Alison L. Cammack: acammac@emory.edu

¹Department of Epidemiology, Rollins School of Public Health, Emory University, 1518 Clifton Road NE, Atlanta, GA 30322, USA

²Department of Psychiatry and Behavioral Sciences, University of Arkansas for Medical Sciences, 4301 W. Markham St., Little Rock, AR 72205, USA

³Department of Pediatrics, University of Arkansas for Medical Sciences, 4301 W. Markham St., Little Rock, AR 72205, USA

⁴Department of Obstetrics and Gynecology, University of Arkansas for Medical Sciences, 4301 W. Markham St., Little Rock, AR 72205, USA

⁵Departments of Psychiatry & Behavioral Sciences and Obstetrics & Gynecology, Leonard M. Miller School of Medicine, University of Miami, 1120 NW 14 Street, Suite 1446, Miami, FL 33136, USA

Abstract

Retrospective reports of exposure to childhood trauma indicate it is common. There is growing interest in relationships between maternal exposure to childhood adversity, perinatal mental health, and pregnancy outcomes. The goal of this study was to describe the self-reported prevalence and test-retest reliability of exposure to childhood maltreatment using the Childhood Trauma Questionnaire among adult women around the time of pregnancy. A substantial proportion of women reported exposure to maltreatment and reliability was generally at least moderate, indicating consistent reporting.

Keywords

Child abuse; Trauma; Pregnancy; Women; Mental health

Introduction

Childhood trauma is commonly reported (CDC 2010). In the Adverse Childhood Experiences study, adult women retrospectively self-reported exposure to sexual (24.7 %), physical (29.0 %), and emotional (13.1 %) abuse and to physical (9.2 %) and emotional (16.7 %) neglect before age 18 (Felitti et al. 1998). Such exposures have been associated

with vulnerability to psychopathology and poor physical health in adults (Shonkoff et al. 2012). Maltreatment is also associated with higher rates of preterm birth and fetal loss (Hillis et al. 2004; Noll et al. 2007) as well as behavioral and psychosocial risk factors for these outcomes among pregnant women, such as substance abuse (Noll et al. 2007) and depression (Chung et al. 2008). With growing interest in preconception health and life course stressors as determinants of perinatal outcomes (Kramer et al. 2011), this literature is likely to expand.

Most studies of the prevalence and long-term implications of child abuse and neglect rely on adult recall of childhood experiences which may be subject to selective recall because of the sensitive/taboo nature of this subject. Studies of test-retest reliability can inform possible bias. The Childhood Trauma Questionnaire (CTQ) is widely used in various populations and has generally shown good test-retest reliability (Bernstein and Fink 1998). However, many of these studies are limited in scope (e.g., utilize short test-retest periods and specialized, non-population-based samples).

To our knowledge, no studies have examined reliability proximate to pregnancy (i.e., preconception, perinatal, and postpartum periods), utilizing the CTQ or any other measure. Studies of the overall adult population may not be generalizable to this context because memory and factors influencing memory such as mood state are influenced by pregnancy itself (Buckwalter et al. 2001). Also, an anecdotal literature suggests that pregnancy-specific events, including having a baby of a certain gender, traumatic deliveries, and initiation of breastfeeding, may influence recall of trauma (Leeners et al. 2006). The aim of this paper is to begin to fill in these gaps by describing women's self-reported exposure to childhood trauma at two times specifically around the time of pregnancy.

Materials and methods

Study participants were drawn from a convenience sample of pregnant women aged 18–45 at the Emory Women's Mental Health Program (WMHP) participating in observational studies of the perinatal course of psychiatric illness from January 2002 to December 2011. Women were referred to the WMHP by community obstetric or psychiatric care providers. Most participants had lifetime histories of a mood and/or anxiety disorder, and many were receiving psychotropic treatment. All participants provided written informed consent, and study participation was independent of their treatment. This research was approved by the Emory University IRB.

Exposures

A retrospective history of childhood trauma was assessed using the Childhood Trauma Questionnaire Short-Form (CTQ; Bernstein and Fink 1998). Three domains of abuse (sexual, physical, and emotional) and two domains of neglect (physical, emotional) are measured with five questions each on a 5-point Likert scale.

Trauma was analyzed both by examining continuous distributions and according to the recommended ordinal cutpoints (no, low, moderate, or severe trauma). Further, all individual trauma scales were dichotomized in two ways: (1) according to the originators'

recommended cutpoints for no or low trauma versus moderate or severe trauma and (2) any versus no trauma. The CTQ also contains a denial scale used to identify possible underreporting of trauma. In the current study, denial was operationalized as any denial at either administration versus no denial.

Women completed the initial CTQ prior to conception or during pregnancy. Administration of the CTQ was repeated after the completion of the pregnancy; however, some women had conceived again and were pregnant or had completed a subsequent pregnancy prior to the second administration of the CTQ. Twelve subjects completed the questionnaire through the mail and the remainder during clinic or research visits.

Analyses

Differences in characteristics of women completing one versus two CTQ assessments were computed with chi-squared tests, Fisher's exact tests, and *t* tests. Test-retest reliability was primarily examined by computing Cohen's *kappas* for dichotomized trauma variables, weighted *kappas* for the four ordinal trauma severity categories, and intraclass correlations (ICCs) for log-transformed continuous measures, all with 95 % confidence intervals. Due to deviations from a 50 % exposure prevalence to individual abuse categories, prevalence-adjusted and bias-adjusted *kappas* (PABAK) were also computed (Byrt et al. 1993). Further, prevalence and bias indices and percent discordance, both overall and differentially according to reports of abuse/lack of abuse (i.e., proportion of positive and negative agreement), were computed as suggested by Byrt et al. (1993). *Kappas* were interpreted according to guidelines from Landis and Koch (0 to 0.2, poor; 0.2 to <0.4, fair; 0.4 to <0.6, moderate; 0.6 to <0.8, substantial; and 0.8 to 1.0, almost perfect) (1977). Regarding ICCs, two-way ANOVA model ICCs were computed to take into account variability that may be associated with preconception/prenatal and postpartum time periods.

Results

A total of 1975 participants completed at least one CTQ; 247 participants completed a second CTQ between 18.1 and 369.9 weeks after the index pregnancy (median time between T1 and T2 =140.7 weeks; IQR 79.3, 213.7). Three women with repeat CTQ administrations were not included in reliability analyses due to administrative timing errors (i.e., second CTQ was administered prior to delivery). All subscales demonstrated good internal consistency (Cronbach's *alpha* at both timepoints .58–.93).

Women were predominantly white, non-Hispanic, and married, and most pregnancies were planned and desired (Table 1). Compared to participants who completed the repeat CTQ, participants completing only one CTQ differed on levels of education, race, marital status, and pregnancy planning and were more likely to report exposure to sexual abuse and physical neglect at the first assessment (sexual abuse 28.31 vs. 22.13 %, $p=.043$; physical neglect 29.21 vs. 20.49 %, $p=.0045$).

For most types of maltreatment, a substantial proportion of participants reported exposure, and inclusion of the low category of abuse substantially increased the reported prevalence for all subtypes except for sexual abuse (Table 2). Log-transformed continuous scores were

slightly higher at T2 than T1 for sexual abuse (mean difference 0.030, 95 % CI .005–.054), but there was no statistically significant change for categorical definitions of abuse. For dichotomous categorizations of maltreatment, McNemar tests revealed no significant differences in the number of women reporting abuse at T1 but not T2 or vice versa. Approximately one quarter of participants ($N=61$) exhibited some denial at either timepoint.

Table 2 shows test-retest reliability for the liberal and conservative dichotomized definitions of the five trauma categories. With the exception of physical neglect (where the *kappa* was below .4 for the more liberal definition), both *kappa* and PABAK scores were at least moderate. PABAK scores tended to be higher than Cohen's *kappa* largely due to the deviation from a 50/50 distribution. The proportion of negative agreement was consistently higher than positive agreement. Weighted *kappas* followed a similar pattern (emotional abuse .67 (.60–.73); physical abuse .67 (.57–.76); sexual abuse .75 (.66–.83); physical neglect .44 (.31–.57); emotional neglect .59 (.51–.67). Intraclass correlations were high for all trauma types (all $r>.7$).

Discussion and conclusion

Retrospective self-reports of childhood trauma were common in this perinatal psychiatric population, with mostly moderate to substantial test-retest reliability between administrations prior to pregnancy outcome and up to 5 years after the pregnancy. Reported prevalences, particularly for the more conservative abuse definition, are similar to those reported by other studies utilizing the CTQ among women (Bernstein and Fink 1998; Matthews et al. 2014). Our findings of generally at least moderate reliability are also similar to those reported elsewhere (Bernstein and Fink 1998; Paivio 2001), although these studies relied on continuous trauma measurements in test-retest assessments.

Regarding the finding of smaller *kappas* for physical neglect, this trauma type is relatively underrepresented in the maltreatment literature (Stoltenborgh et al. 2013), as compared to other forms of maltreatment. In our study population, as likely others, women were more likely to be screened for other traumas as part of clinical assessments, which in our case usually occurred before administration of the CTQ. Thus, it is possible differential response priming may have influenced reliability.

Utilizing a study population that was recruited from a population of relatively homogeneous upper middle class white women with a history of psychiatric illness who participated in longitudinal studies introduces two noteworthy limitations. First, these highly compliant and motivated women may have been more likely to report abuse and neglect. Second, relatively homogeneous populations may yield decreased estimates of reliability since increased precision is needed to distinguish differences; therefore, compared to more diverse populations of women around the time of pregnancy, our observed effect sizes may be underestimates.

In conclusion, this study offers some reassurance that the CTQ administered around the time of pregnancy yields adequate consistency. However, reliability does not necessarily translate to validity. Further research of misclassification of maltreatment as well as the magnitude and direction of any potential biases, particularly as they relate to associations between these

exposures and perinatal outcomes, remains important. Related studies, including those utilizing alternative measures of trauma and those assessing the prevalence and reliability of other types of childhood adversity (e.g., poverty, parental conflict), should also be conducted.

Acknowledgments

ALC received support from the National Institutes of Health training grant T32HD052460 and Health Resources and Services Administration Maternal Child Health Epidemiology training grant T03MC07651.

References

- Bernstein DP, Fink L. Childhood trauma questionnaire: a retrospective self-report. Psychological, San Antonio. 1998
- Buckwalter JG, Buckwalter DK, Bluestein BW, Stanczyk FZ. Pregnancy and post partum: changes in cognition and mood. *Prog Brain Res*. 2001; 133:303–319.10.1016/S0079-6123(01)33023-6 [PubMed: 11589139]
- Byrt T, Bishop J, Carlin JB. Bias, prevalence and kappa. *J Clin Epidemiol*. 1993; 46(5):423–429.10.1016/0895-4356(93)90018-V [PubMed: 8501467]
- Centers for Disease Control and Prevention (CDC). Adverse childhood experiences reported by adults —five states, 2009. *MMWR Morb Mortal Wkly Rep*. 2010; 59(49):1609–1613. [PubMed: 21160456]
- Chung EK, Mathew L, Elo IT, Coyne JC, Culhane JF. Depressive symptoms in disadvantaged women receiving prenatal care: the influence of adverse and positive childhood experiences. *Ambul Pediatr*. 2008; 8(2):109–116.10.1016/j.ambp.2007.12.003 [PubMed: 18355740]
- Felitti VJ, Anda RF, Nordenberg D, Williamson DF, Spitz AM, Edwards V, Koss MP, Marks JS. Relationship of childhood abuse and household dysfunction to many of the leading causes of death in adults. The Adverse Childhood Experiences (ACE) Study. *Am J Prev Med*. 1998; 14(4):245–258.10.1016/S0749-3797(98)00017-8 [PubMed: 9635069]
- Hillis SD, Anda RF, Dube SR, et al. The association between adverse childhood experiences and adolescent pregnancy, long-term psychosocial consequences, and fetal death. *Pediatrics*. 2004; 113:320–327.10.1542/peds.113.2.320 [PubMed: 14754944]
- Kramer MR, Hogue CJ, Dunlop AL, Menon R. Preconceptional stress and racial disparities in preterm birth: an overview. *Acta Obstet Gynecol Scand*. 2011; 90(12):1307–1316.10.1111/j.1600-0412.2011.01136.x [PubMed: 21446927]
- Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977; 33(1):159–174.10.2307/2529310 [PubMed: 843571]
- Leeners B, Richter-Appelt H, Imthurn B, Rath W. Influence of childhood sexual abuse on pregnancy, delivery, and the early post-partum period in adult women. *J Psychosom Res*. 2006; 61(2):139–151.10.1016/j.jpsychores.2005.11.006 [PubMed: 16880016]
- Matthews KA, Chang YF, Thurston RC, Bromberger JT. Child abuse is related to inflammation in mid-life women: role of obesity. *Brain Behav Immun*. 2014; 36:29–34.10.1016/j.bbi.2013.09.013 [PubMed: 24076375]
- Noll JG, Schulkin J, Trickett PK, Susman EJ, Breech L, Putnam FW. Differential pathways to preterm delivery for sexually abused and comparison women. *J Pediatr Psychol*. 2007; 32(10):1238–1248.10.1093/jpepsy/jsm046 [PubMed: 17569710]
- Paivio SC. Stability of retrospective self-reports of child abuse and neglect before and after therapy for child abuse issues. *Child Abuse Negl*. 2001; 25(8):1053–1068.10.1016/S0145-2134(01)00256-3 [PubMed: 11601597]
- Shonkoff JP, Garner AS. Committee on Psychosocial Aspects of Child and Family Health; Committee on Early Childhood, Adoption, and Dependent Care. The lifelong effects of early childhood adversity and toxic stress. *Pediatrics*. 2012; 129(1):e232–e246.10.1542/peds.2011-2663 [PubMed: 22201156]

Stoltenborgh M, Bakermans-Kranenburg MJ, van IJzendoorn MH. The neglect of child neglect: a meta-analytic review of the prevalence of neglect. *Soc Psychiatry Psychiatr Epidemiol.* 2013; 48(3):345–355.10.1007/s00127-012-0549-y [PubMed: 22797133]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1

Characteristics of women completing one or two CTQ questionnaires

	<u>Women completing two CTQs</u>		<u>Women completing one CTQ</u>	
	<i>N</i> Mean=32.7 years	% SD=4.5	<i>N</i> Mean=33.3 years	% SD=5.5
Age at delivery				
Education ***				
12 years or less	2	0.82	151	8.90
13–15	42	17.21	360	21.21
16+	200	81.97	1186	69.89
Race **				
Asian	8	3.28	34	1.96
Black	12	4.92	170	9.82
Native American	0	0.00	23	1.33
Multiple races	3	1.23	30	1.73
Unknown	1	0.41	19	1.10
Pacific islander	0	0.00	1	0.06
White	220	90.16	1454	84
Ethnicity				
Hispanic	5	2.06	65	3.76
non-Hispanic	239	97.94	1666	96.24
Marital status ***				
Divorced	5	2.06	58	3.36
Married	217	88.93	1402	81.18
Never married, lives with partner	7	2.87	158	9.15
Never married, lives alone	11	4.51	90	5.21
Separated	3	1.23	14	0.81
Unknown	1	0.41	1	0.06
Widowed	0	0.00	4	0.23
Planned pregnancy *				
No	54	22.41	455	28.78
Yes	187	77.59	1126	71.22
Pregnancy wantedness				
No	5	2.09	67	4.34
Ambivalent	34	14.23	201	13.01
Yes	200	83.68	1277	82.65
SCID diagnoses				
Major depression	152	63.07	461	62.13
Depression NOS	4	1.66	8	1.08
Bipolar				
Bipolar I	45	18.91	139	19.04
Bipolar 2	10	4.20	25	3.24
GAD	66	27.39	168	22.64

Age at delivery	<u>Women completing two CTQs</u>		<u>Women completing one CTQ</u>	
	<i>N</i> Mean=32.7 years	% SD=4.5	<i>N</i> Mean=33.3 years	% SD=5.5
PTSD	42	17.43	161	21.70
Alcohol dependence	87	36.10	275	37.06
Eating disorders				
Anorexia	14	5.81	41	5.53
Bulimia	11	4.56	51	6.57
Binge eating	11	4.72	51	7.23

* $p < .05$ for difference between participants with one versus two CTQ administrations using chi-square test;

** $p < .05$ for difference between participants with one versus two CTQ administrations, using Fisher's exact test;

*** $p < .01$ for difference between participants with one versus two CTQ administrations, using Fisher's exact test

Table 2

Prevalence and reliability of individual traumas

Trauma subtype	Kappa		PABAK	Observed Agreement (Po)	Expected Agreement (Pe)	Positive Agreement (Ppos)	Negative Agreement (Pneg)	Prevalence Index	Bias Index
Emotional abuse (none versus any)									
2nd CTQ									
1st CTQ	N	Y	0.75 (0.66–0.83)	0.88	0.51	0.86	0.89	−0.14	0.01
N	123	14							
Y	16	90							
43.62 %									
42.80 % 243									
Physical abuse (none versus any)									
2nd CTQ									
1st CTQ	N	Y	0.62 (0.50–0.73)	0.86	0.64	0.71	0.91	−0.52	−0.01
N	169	18							
Y	16	41							
23.36 %									
24.18 % 244									
Sexual abuse (none versus any)									
2nd CTQ									
1st CTQ	N	Y	0.76 (0.67–0.86)	0.91	0.63	0.82	0.94	−0.52	−0.04
N	175	15							
Y	6	48							
22.13 %									
25.82 % 244									
Physical neglect (none versus any)									
2nd CTQ									
1st CTQ	N	Y	0.37 (0.23–0.51)	0.79	0.66	0.50	0.87	−0.57	−0.02
N	165	28							
Y	23	26							
20.25 %									
22.31 % 242									
Emotional neglect (none versus any)									
2nd CTQ									
1st CTQ	N	Y	0.60 (0.50–0.70)	0.80	0.50	0.78	0.82	−0.09	0.04
N	108	20							

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Trauma subtype	Y	29	87	47.54 %	Kappa	PABAK	Observed Agreement (Po)	Expected Agreement (Pe)	Positive Agreement (Ppos)	Negative Agreement (Pneg)	Prevalence Index	Bias Index
Emotional abuse (none or low versus moderate or severe)												
2nd CTQ												
1st CTQ	N	Y			0.62 (0.49–0.74)	0.75	0.88	0.68	0.69	0.92	−0.60	−0.01
N	179	16										
Y	14	34		19.75 %								
			20.58 %	243								
Physical abuse (none or low versus moderate or severe)												
2nd CTQ												
1st CTQ	N	Y			0.65 (0.51–0.80)	0.84	0.92	0.77	0.70	0.96	−0.74	−0.03
N	203	13										
Y	6	22		11.48 %								
			14.34 %	244								
Sexual abuse (none or low versus moderate or severe)												
2nd CTQ												
1st CTQ	N	Y			0.73 (0.62–0.84)	0.84	0.92	0.71	0.78	0.95	−0.65	−0.03
N	192	13										
Y	6	33		15.98 %								
			18.85 %	244								
Physical neglect (none or low versus moderate or severe)												
2nd CTQ												
1st CTQ	N	Y			0.52 (0.34–0.70)	0.83	0.92	0.83	0.57	0.95	−0.81	−0.02
N	209	12										
Y	8	13		8.68 %								
			10.33 %	242								
Emotional neglect (none or low versus moderate or severe)												
2nd CTQ												
1st CTQ	N	Y			0.61 (0.48–0.74)	0.76	0.88	0.70	0.68	0.93	−0.63	−0.01
N	184	16										
Y	13	31		11.48 %								

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Trauma subtype	Kappa	PABAK	Observed Agreement (Po)	Expected Agreement (Pe)	Positive Agreement (Ppos)	Negative Agreement (Pneg)	Prevalence Index	Bias Index
	14.34 %		244					