Online supplement
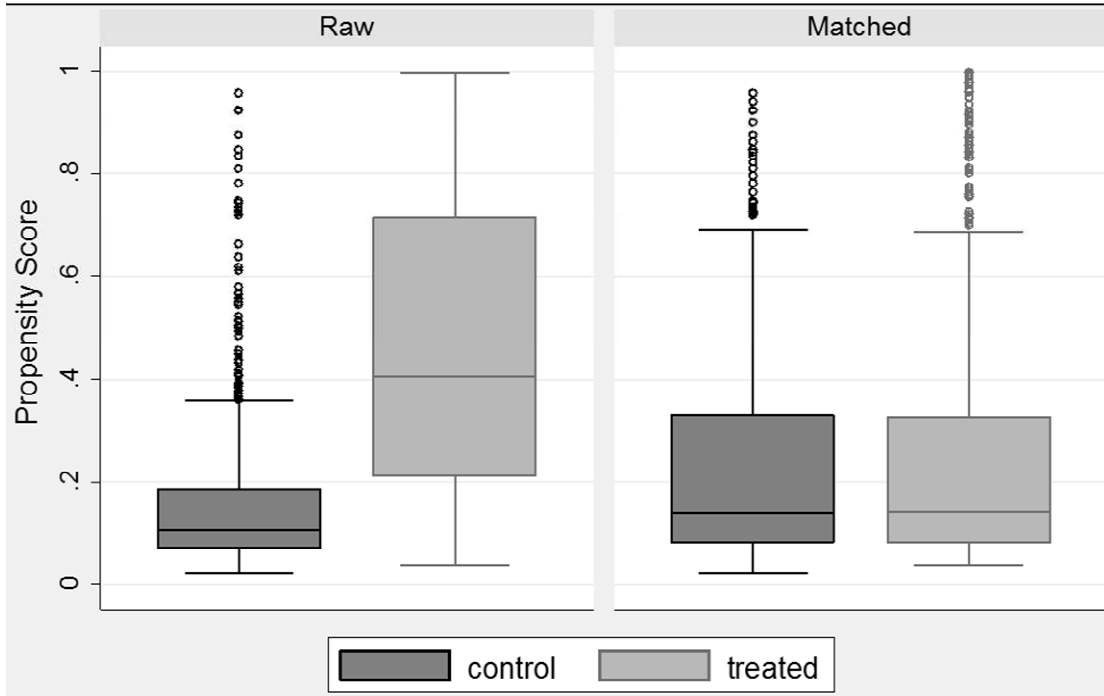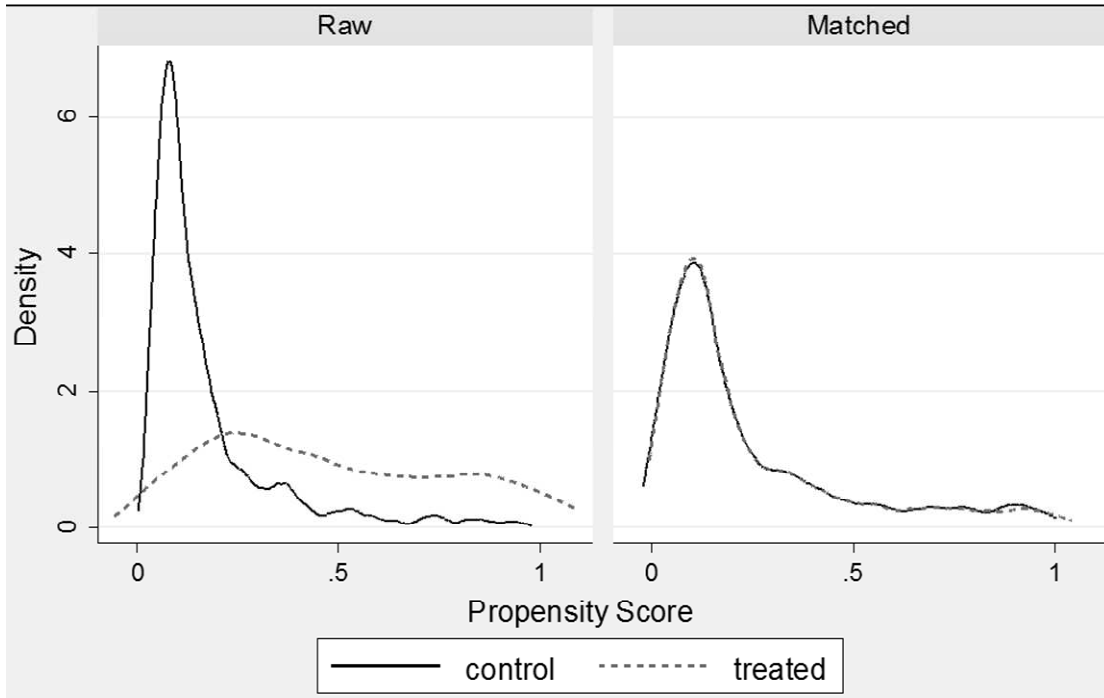
Baseline and follow-up differences among eligible analytic sample, with comparisons separated by reason for non-treatment					
	Comparison				Treatment
	Opted out (n=139)		EMERALD not available at site (n=360)		EMERALD (n=162)
	Mean ± Std. Dev.	%	Mean ± Std. Dev.	%	Mean ± Std. Dev.	%
Female		67%		69%		81%
white non-Hispanic		71%		86%		72%
Age	15.0 ± 1.7		14.8 ± 1.6		14.9 ± 1.5	
Baseline PHQ-9A score	13.7 ± 4.2		14.7 ± 4.2		15.2 ± 4.1	
Number of measurements in extracted data	5.5 ± 5.1		5.2 ± 4.3		13.5 ± 8.3	
Number of days from baseline to latest measure in data extraction	254.8 ± 304.5		212.1 ± 233.7		373.9 ± 298.3	

Density and box plots indicating balance of propensity scores in raw versus matched data

Full results from logistic regression of treatment response and remission of depression on EMERALD participation and covariates								
	Treatment response				Depression remission			
	Odds ratio	p-value	95% Conf. Int.		Odds ratio	p-value	95% Conf. Int.	
In EMERALD	2.03	0.003	1.28	3.23	1.86	0.019	1.10	3.12
Age	1.01	0.863	0.91	1.13	1.08	0.197	0.96	1.23
Female	0.75	0.156	0.51	1.11	0.64	0.046	0.41	0.99
white, non-Hispanic	1.17	0.510	0.73	1.88	1.19	0.535	0.69	2.05
Number of measurements in extracted data	1.07	0.007	1.02	1.11	1.07	0.003	1.02	1.12
Number of days from baseline to latest measure in data extraction	1.00	0.046	1.00	1.00	1.00	0.072	1.00	1.00
Had six-month follow-up (did not use last observation carried forward)	2.32	<0.001	1.52	3.55	2.27	0.001	1.42	3.64
Baseline PHQ-9A score	1.06	0.004	1.02	1.11	0.96	0.163	0.91	1.02
Use of interpreter	2.12	0.433	0.32	13.91	1.91	0.468	0.33	11.05
Intercept	0.08	0.004	0.01	0.45	0.07	0.009	0.01	0.51