

Table 2 Estimates of the impact of different definitions of improved water supply on the prevalence of diarrhoea in children under 5 years old: comparison of propensity score matching (PSM), coarsened exact matching (CEM) and logistic regression analysis

Definition/analysis method	No. of treated	No. of controls	Estimate (% point change) ^a	SE of % point change	95% CI of % point change	P-value
Main definitions						
<i>Improved-uninterrupted definition:</i>						
<i>Improved-uninterrupted vs improved-interrupted</i>						
PSM	5634	3877	-0.026	0.009	-0.043 to -0.009	< 0.01
CEM	3356	2501	-0.019	0.006	-0.029 to -0.007	< 0.001
Logistic regression	5738	3993	-0.031	0.005	-0.042 to -0.021	< 0.001
<i>Improved-unstored definition:</i>						
<i>Improved-unstored vs improved-stored</i>						
PSM	7764	1779	-0.035	0.012	-0.059 to -0.012	< 0.01
CEM	3289	1259	-0.027	0.007	-0.040 to -0.013	< 0.001
Logistic regression	7909	1822	-0.030	0.006	-0.044 to -0.016	< 0.001
Combined effects analysis						
<i>Combined effects analysis 1:</i>						
<i>Improved-uninterrupted-unstored vs all other improved</i>						
PSM	5302	4195	-0.027	0.009	-0.043 to -0.010	< 0.01
CEM	3232	2654	-0.020	0.006	-0.030 to -0.008	< 0.001
Logistic regression	5395	4336	-0.032	0.005	-0.042 to -0.022	< 0.001
<i>Combined effects analysis 2:</i>						
<i>Improved-uninterrupted-unstored vs improved-interrupted-stored</i>						
PSM	5302	1409	-0.050	0.015	-0.079 to -0.022	< 0.001
CEM	1996	868	-0.030	0.008	-0.045 to -0.015	< 0.001
Logistic regression	5395	1479	-0.047	0.006	-0.064 to -0.029	< 0.001
<i>Combined effects analysis 3:</i>						
<i>Improved-uninterrupted-unstored vs improved-interrupted-unstored</i>						
PSM	5302	2443	-0.028	0.009	-0.046 to -0.010	< 0.01
CEM	2721	1574	-0.013	0.006	-0.025 to -0.001	< 0.05
Logistic regression	5395	2514	-0.028	0.005	-0.040 to -0.015	< 0.001
Rural-urban analysis						
<i>Improved-uninterrupted vs improved-interrupted: children residing in rural households only</i>						
PSM	3372	2684	-0.047	0.010	-0.068 to -0.027	< 0.001
CEM	2130	1818	-0.025	0.009	-0.043 to -0.007	< 0.01
Logistic regression	3442	2748	-0.036	0.006	-0.049 to -0.023	< 0.001
<i>Improved-uninterrupted vs improved-interrupted: children residing in urban households only</i>						
PSM	2260	1186	-0.017	0.014	-0.045 to 0.011	
CEM	1286	730	-0.008	0.011	-0.030 to 0.013	
Logistic regression	2296	1245	-0.022	0.009	-0.041 to -0.004	< 0.05

^aEstimates for the CEM and logit models are marginal effects
SE = standard error; CI = confidence interval.