

Supplementary Table S1. Logistic regression results for exposure–outcome associations (N = 70)

Comparison	Exposed n/N (%)	Unexposed n/N (%)	Unadjusted OR (95% CI)	p	Adjusted OR* (95% CI)	p
Symptom flare-up: always-untreated vs never	29/29 (100%)	3/27 (11.1%)	413.0 (20.3–8389.5) [†]	<0.001	Not estimable (quasi-complete separation)	—
Symptom flare-up: Lake Sonfon vs other main source	35/43 (81.4%)	3/27 (11.1%)	35.0 (8.4–145.5)	<0.001	44.1 (8.6–227.5)	<0.001
Respiratory illness: always-untreated vs never	10/29 (34.5%)	2/27 (7.4%)	6.6 (1.3–33.6)	0.021	24.2 (2.5–235.3)	0.006
Respiratory illness: Lake Sonfon vs other main source	10/43 (23.3%)	2/27 (7.4%)	3.8 (0.8–18.9)	0.111	8.1 (1.1–59.1)	0.038

**Adjusted for occupation (farmer/miner/trader) by multivariable logistic regression. [†]Haldane-Anscombe correction (0.5 added to each cell of the 2×2 table) applied because of a zero cell. Exposed/unexposed group definitions: rows 1 and 3 compare households reporting "always" using untreated lake water (n = 29) against those reporting "never" (n = 27), excluding the intermediate "sometimes" category (n = 14); rows 2 and 4 compare households reporting Lake Sonfon as their main drinking-water source (n = 43) against those using any other source (n = 27). p-values for unadjusted estimates are two-sided Fisher's exact tests; p-values for adjusted estimates are Wald tests from the logistic regression model.*