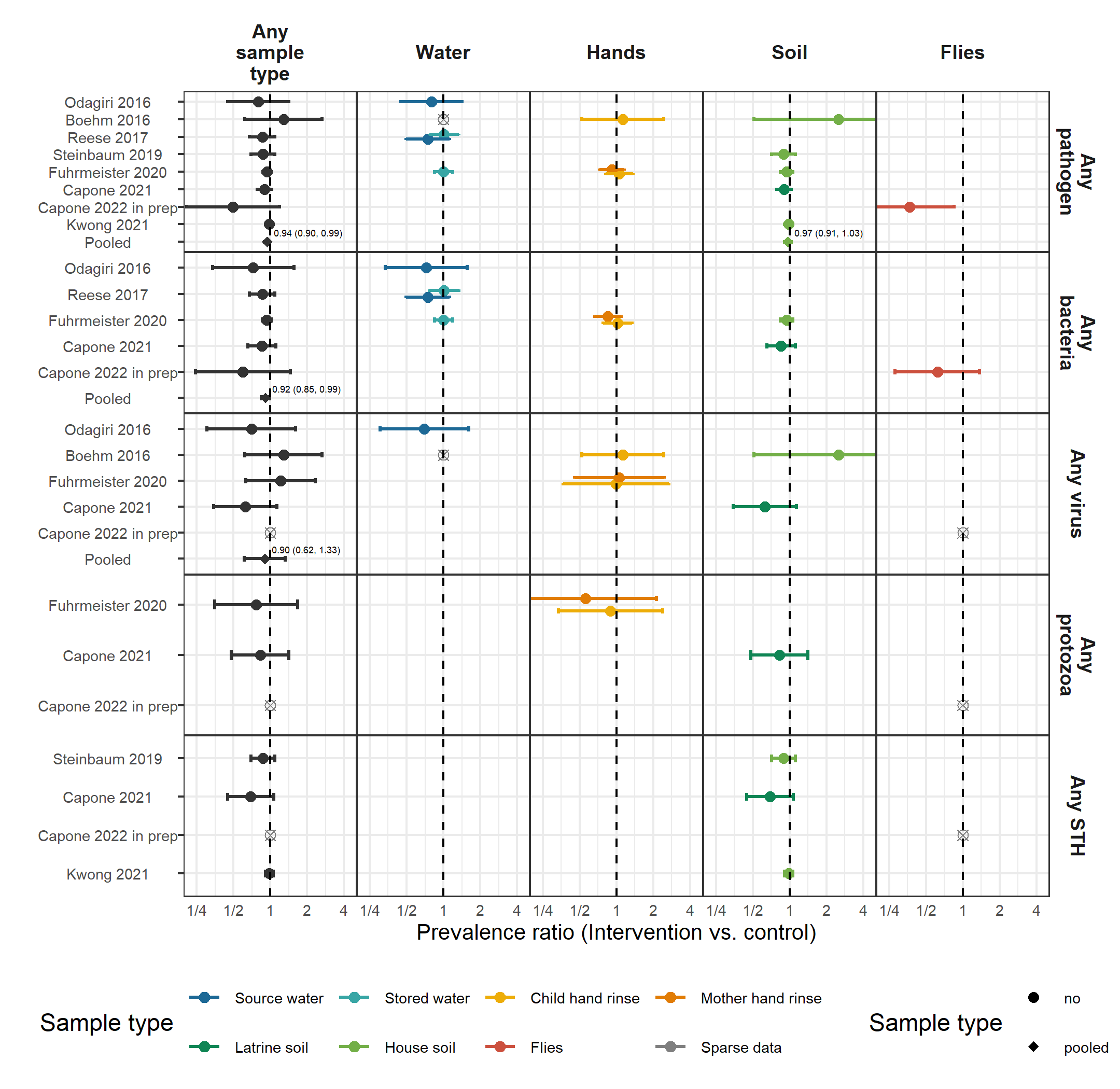
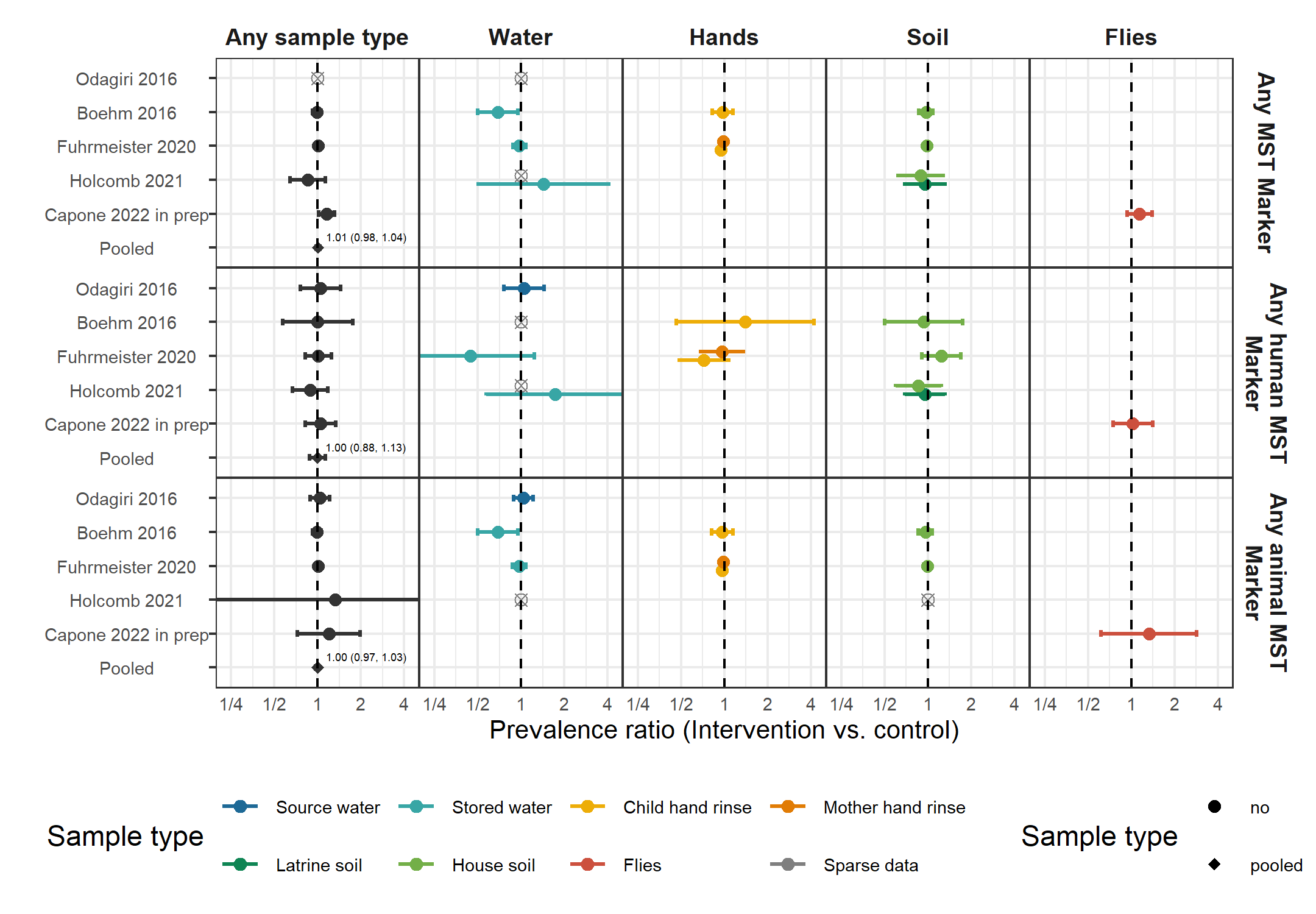
Effects of water, sanitation, and hygiene interventions on detection of enteropathogens and host-specific faecal markers in the environment: an individual-participant data meta-analysis

Andrew Mertens, Benjamin F. Arnold, Jade Benjamin-Chung, Alexandria Boehm, Joe Brown, Drew Capone, Thomas Clasen, Erica Fuhrmeister, Jessica Grembi, David Holcomb, Jackie Knee, Laura H Kwong, Audrie Lin, Stephen P. Luby, Rassul Nala, Kara Nelson, Sammy Njenga, Clair Null, Amy J. Pickering, Mahbubur Rahman, Heather Reese, Lauren Steinbaum, Jill Stewart, Ruwan Thilakaratne, Oliver Cumming, John M. Colford Jr., Ayse Ercumen

## Figures



**Figure 1.** Forest plots of WASH intervention effects on the prevalence of any enteropathogen or type of enteropathogen (any bacteria, any virus, any protozoa and any STH) in different types of environmental samples. Pooled estimates are presented when there are four or more study-specific estimates for a specific sample type and target combination and are denoted with diamond-shaped points. Grey crossed points denote data that were too sparse to estimate a prevalence ratio (i.e., <10 positive observations). Samples of the same type from different locations (source vs. stored water, flies in kitchen vs. latrine, soil from courtyard vs. latrine) or different individuals (child vs. mother’s hands) are plotted separately. Point estimates and confidence intervals are printed next to pooled estimates. All estimates are adjusted for potential confounders.



**Figure 2.** Forest plots of WASH intervention effects on the prevalence of any MST marker or type of MST marker (human or animal MST markers) in different types of environmental samples. Pooled estimates are presented when there are four or more study-specific estimates for a specific sample type and target combination and are denoted with diamond-shaped points. Grey crossed points denote data that were too sparse to estimate a prevalence ratio (i.e., <10 positive observations). Samples of the same type from different locations (source vs. stored water, flies in kitchen vs. latrine, soil from courtyard vs. latrine) or different individuals (child vs. mother’s hands) are plotted separately. Point estimates and confidence intervals are printed next to pooled estimates. All estimates are adjusted for potential confounders.

## Tables

### Table 1. Characteristics of included publications

| **Parent study** | **Study design** | **Intervention** | **Time between intervention and environmental sampling** | **Location** | **Author/ year** | **Sample types** | **Targets** | **Analytic method** | **Number of samples** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| WASH Benefits Bangladesh | Cluster-randomized trial | Latrine upgrades, child potties, scoops for feces disposal | 4 months | Rural Bangladesh | Boehm et al. 2016 | Stored drinking water, child hands, soil | Rotavirus, General, human, avian and ruminant fecal markers | qPCR | 1,482 |
| - | - | - | 16-35 months | - | Fuhrmeister et al. 2020 | Stored drinking water, child and mother hands, soil | Pathogenic E. coli, norovirus, Giardia | qPCR | 2,601 |
| - | - | - | ~2 years | - | Kwong et al. 2021 | Courtyard soil | Soil-transmitted helminths | Microscopy | 1,396 |
| WASH Benefits Kenya | Cluster-randomized trial | Latrine upgrades, child potties, scoops for feces disposal | ~2 years | Rural Kenya | Steinbaum et al. 2019 | Courtyard soil | Soil-transmitted helminths | Microscopy | 2,149 |
| MapSan | Controlled before-and-after study | Latrine upgrades | ~1 year | Urban Mozambique | Holcomb et al. 2020 | Source and stored water, household and latrine soil, food | General, human and avian fecal MST markers | qPCR | 353 |
| - | - | - | ~1 year | - | Capone et al. 2021 | Household and latrine soil | Panel of 18 enteric pathogens | qPCR | 88 |
| - | - | - | ~2 years | - | Capone et al. 2022 in prep. | Flies caught in latrine and kitchen | Panel of 16 enteric pathogens and MST markers | qPCR | 86 |
| Gram Vikas | Matched cohort study | Latrine upgrades, piped water | ~6-10 years | Rural India | Reese et al. 2017 | Source and stored water | V. cholerae, Shigella | Slide agglutination serotyping | 3,452 |
| Total Sanitation Campaign | Cluster-randomized trial | Latrine upgrades | ~1 year | Rural India | Odagiri et al. 2016 | Source water | V. cholerae, rotavirus, adenovirus,general, human, and animal fecal markers | qPCR, microscopy | 60 |

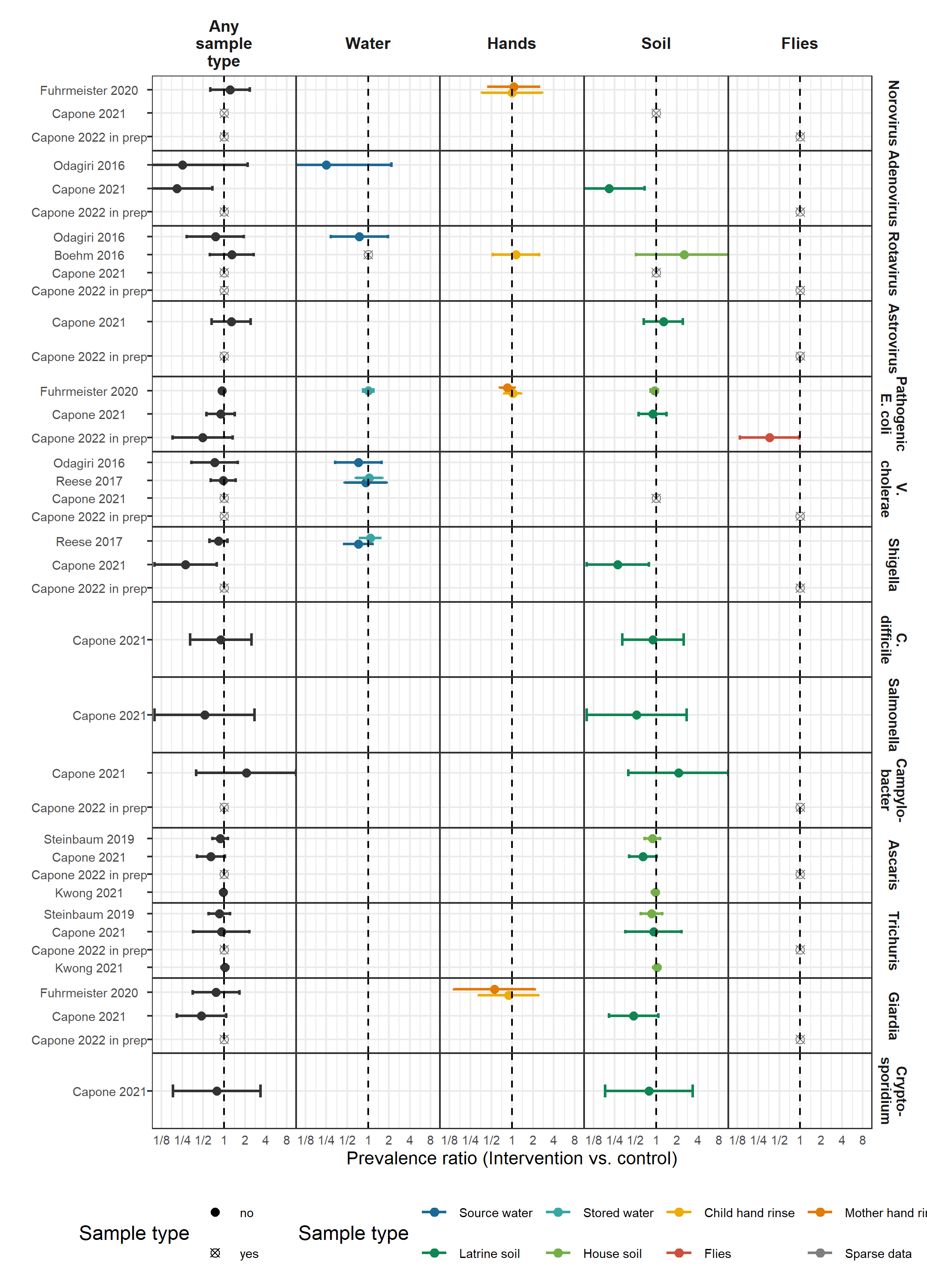
### Table 2. Mean (SD) abundances of enteropathogen and MST targets by study arm. Means are log10-transformed gene copies for MST markers and mean egg counts for soil transmitted helminths (*Ascaris* and *Trichuris*). Intervention effects are shown as adjusted differences in log10-transformed gene copies and ratios of helminth egg counts between the intervention and control arms.

| **Study** | **Sample** | **Target** | **N** | **% in ROQ** | **Control mean, median (SD)** | **Intervention mean, median (SD)** | **Intervention effect (95% CI)** | **P value** | **Wilcoxon P value** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Fuhrmeister 2020 | Child hand rinse | Animal (BacCow) | 365 | 75.9 | 3.6, 3.9 (1.4) | 3.4, 3.8 (1.4) | -0.17 (-0.47 0.12) | 0.25 | 0.17 |
| - | Mother's hand rinse | Animal (BacCow) | 725 | 66.5 | 3.3, 3.8 (1.4) | 3, 3.7 (1.5) | -0.28 (-0.49 -0.07) | 0.01 | 0.01 |
| Holcomb 2021 | Latrine soil | Human (M. smithii) | 113 | 51.3 | 6.7, 6.5 (0.6) | 6.5, 6.3 (0.5) | -0.14 (-0.38 0.11) | 0.27 | 0.58 |
| Capone 2022 in prep |  | Human (BacHum) | 173 | 77.5 | 3.8, 3.8 (1.3) | 4, 4.2 (0.9) | 0.14 (-0.19 0.47) | 0.41 | 0.07 |
| Steinbaum 2019 | House soil | Ascaris | 2,101 | 100.0 | 2.2, 0 (18.8) | 1.4, 0 (9.3) | 0.65 (0.33 1.28)a | 0.21 | 0.33 |
| - | - | Trichuris | 2,102 | 100.0 | 0.2, 0 (1.8) | 0.2, 0 (1) | 0.73 (0.36 1.48)a | 0.38 | 0.39 |
| Kwong 2021 | House soil | Ascaris | 1,426 | 100.0 | 2.3, 0.7 (6.7) | 2.2, 0.6 (6.9) | 0.97 (0.68 1.38)a | 0.85 | 0.54 |
| - | - | Trichuris | 1,426 | 100.0 | 1.6, 0.4 (5) | 2, 0.4 (5) | 1.22 (0.87 1.71)a | 0.26 | 0.17 |

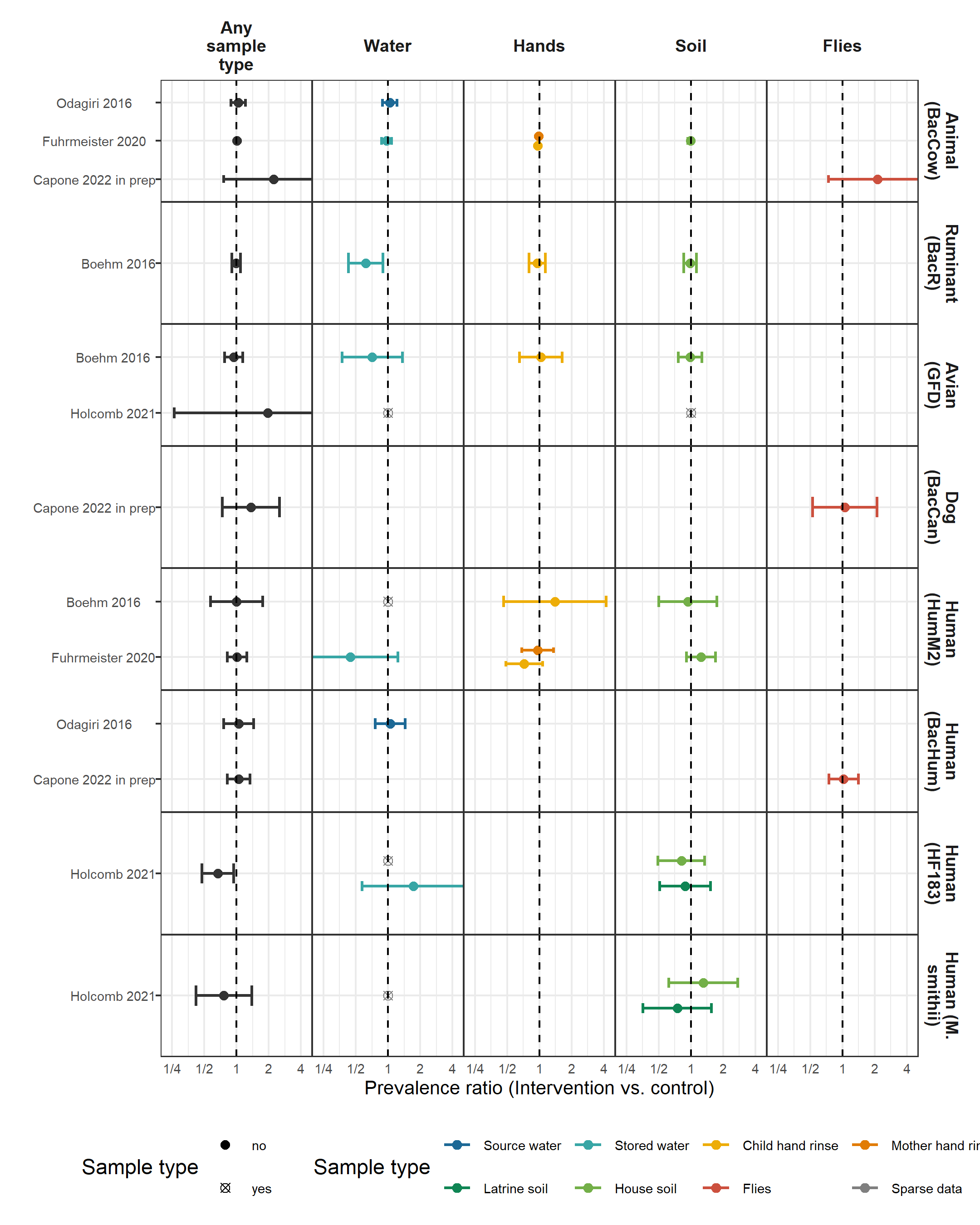
ROQ: Range of quantification; SD: Standard deviation; CI: Confidence interval; Wilcoxon P-value: Non-parametric Wilcoxon rank sum test P-value.

a Marks ratio estimates from negative binomial models.

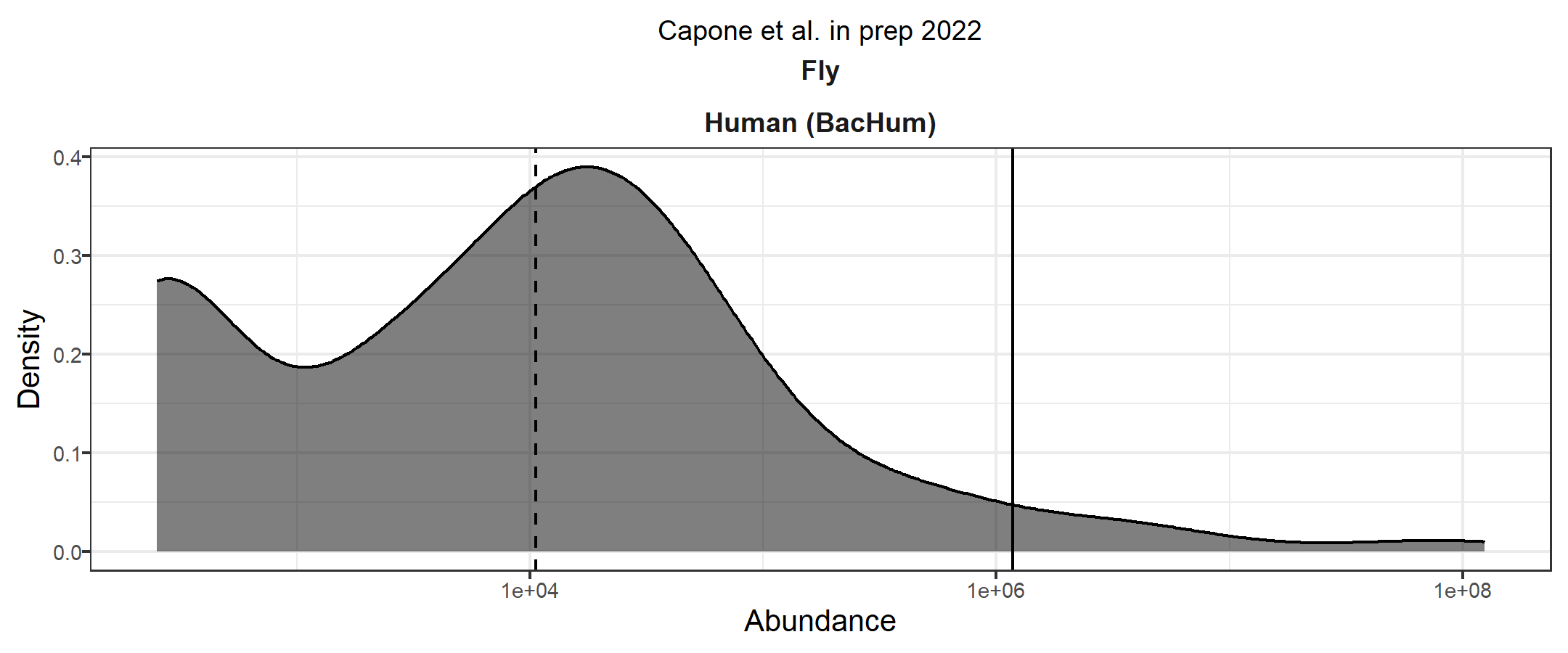
## Supplementary Figures

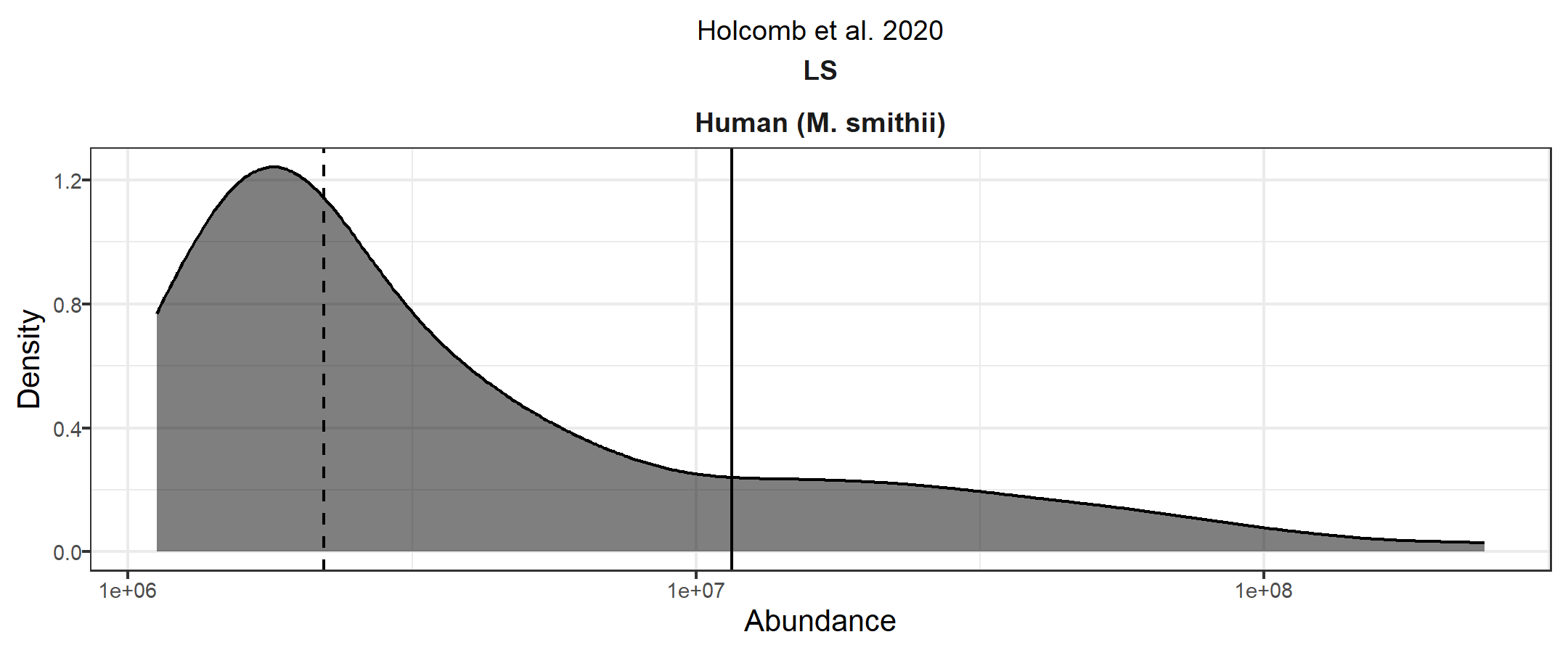


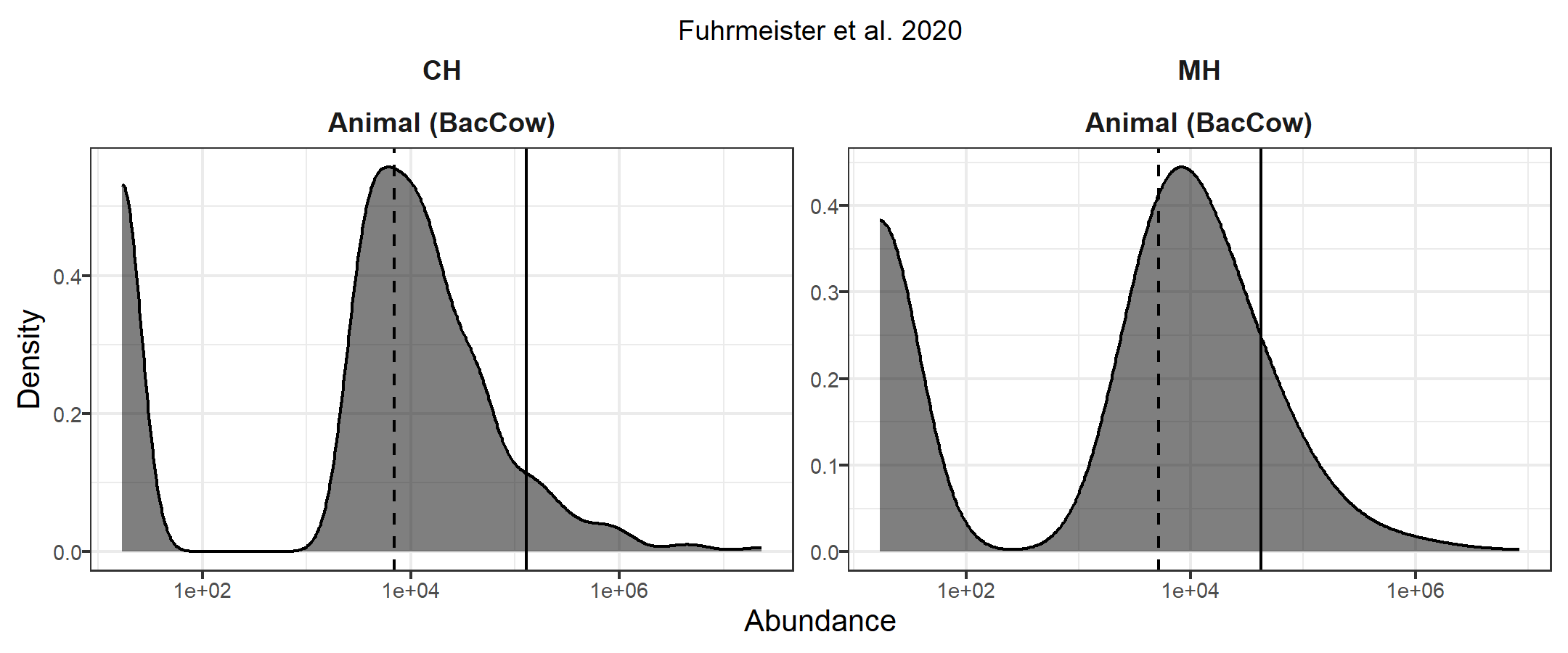
**Figure S2.** Forest plots of intervention effects on the prevalence of specific pathogens.

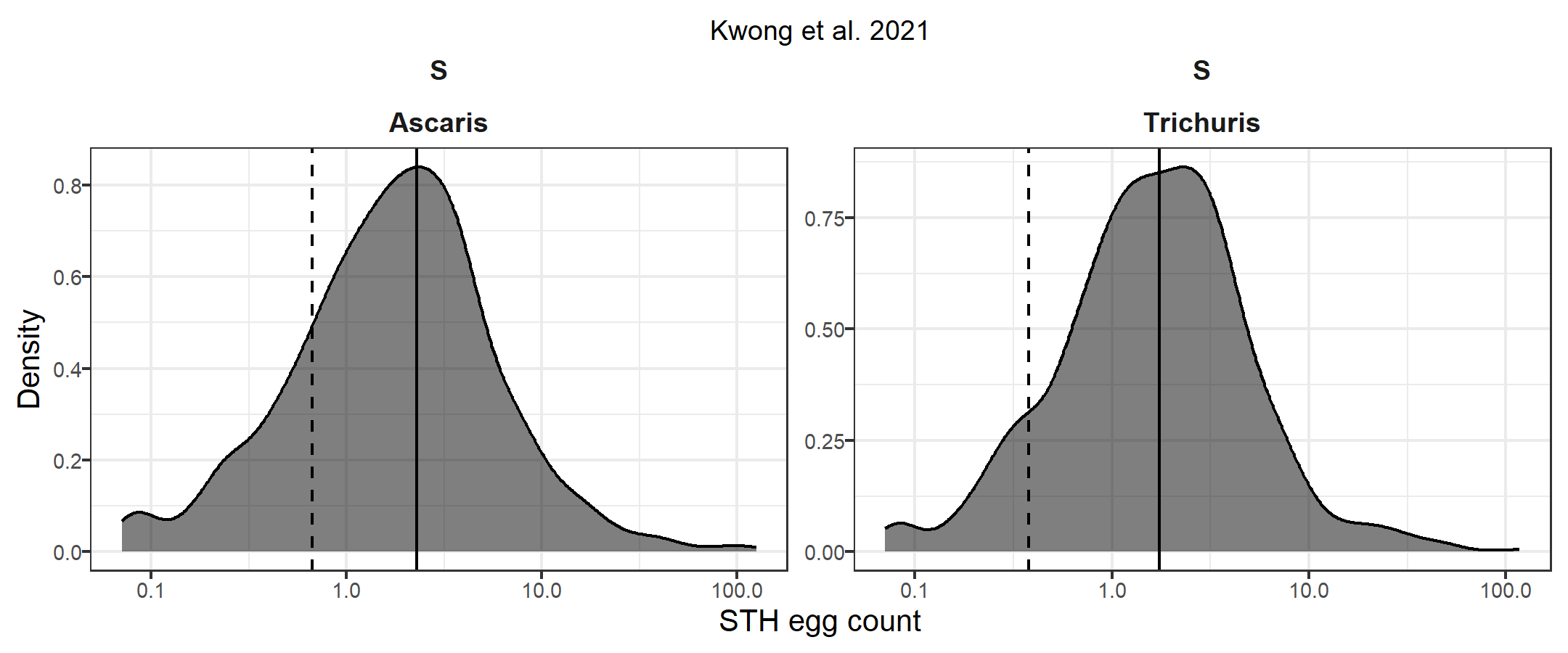


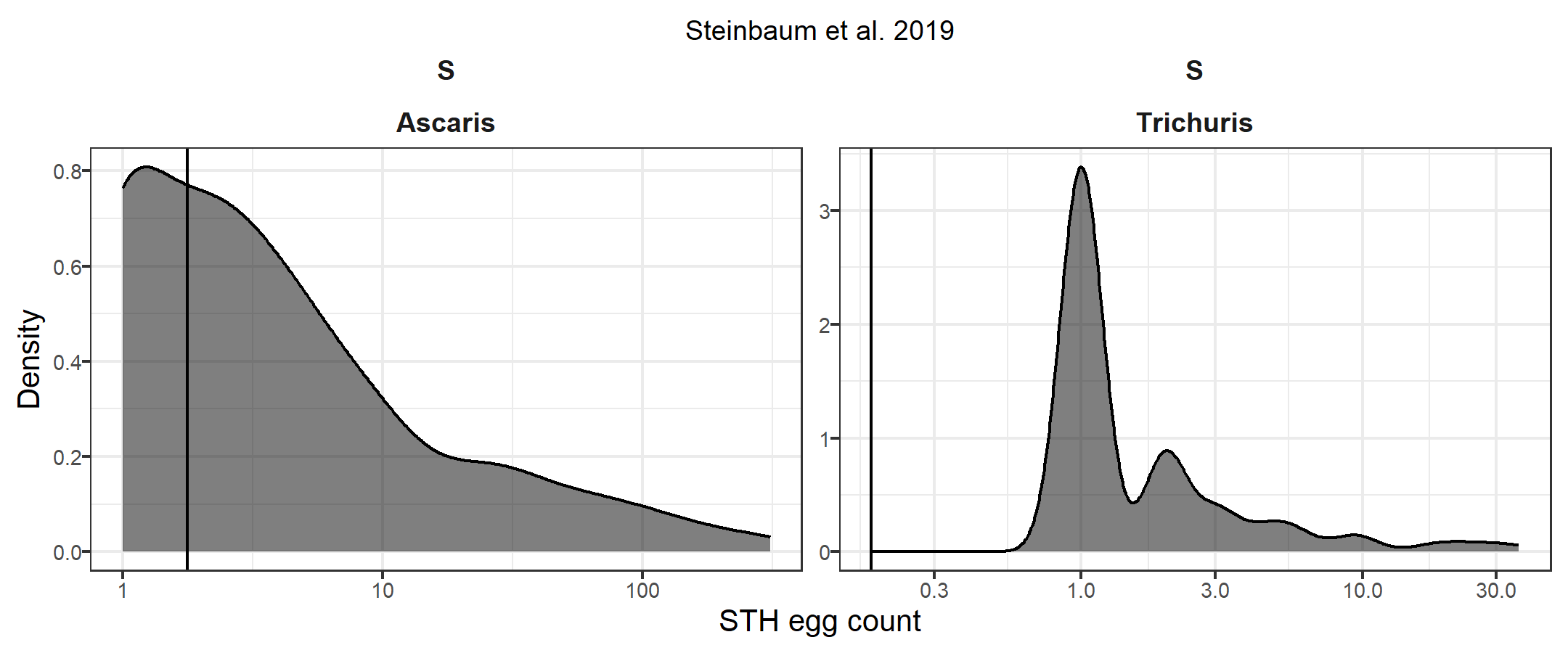
**Figure S3.** Forest plots of intervention effects on the prevalence of specific MST markers.



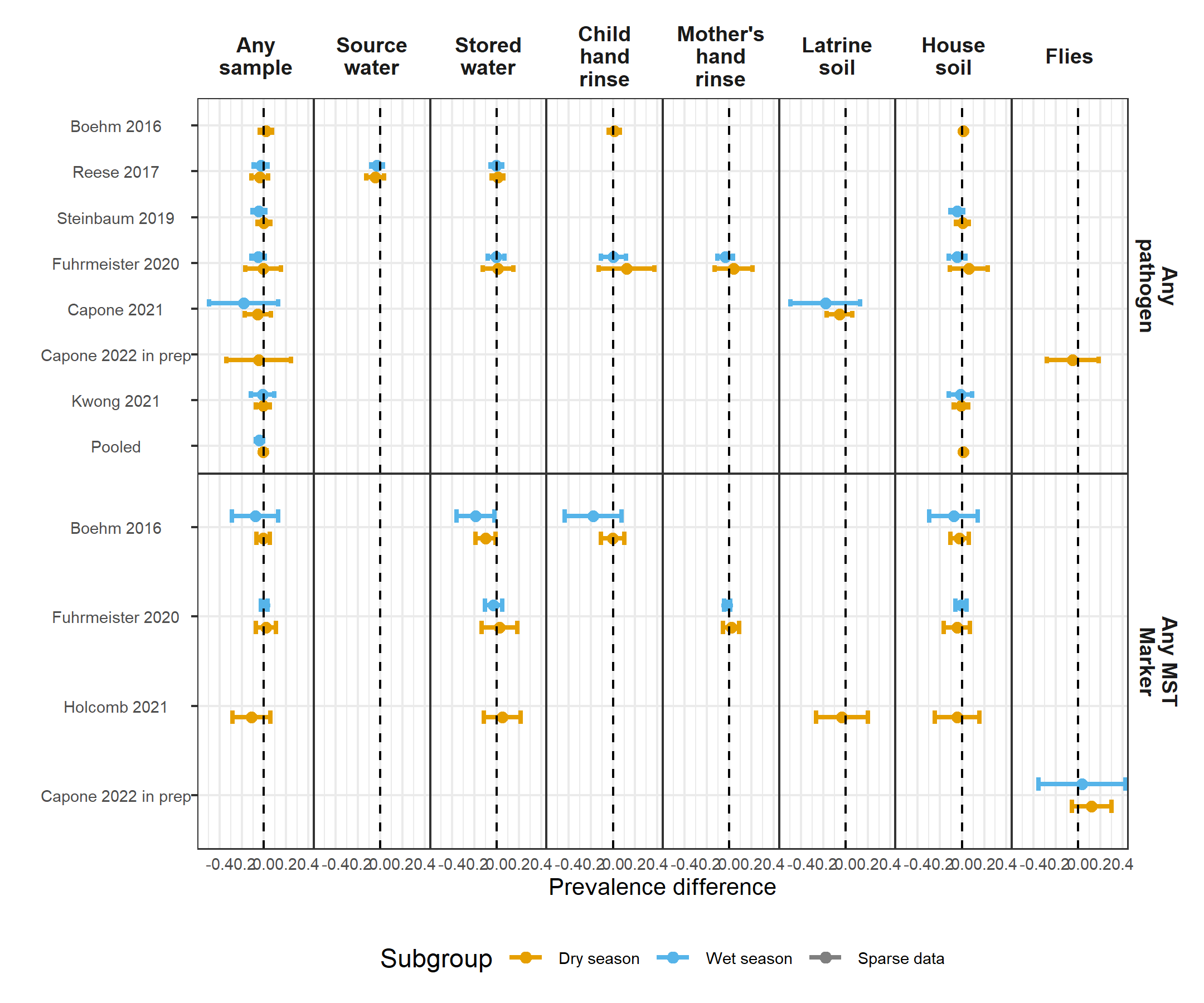






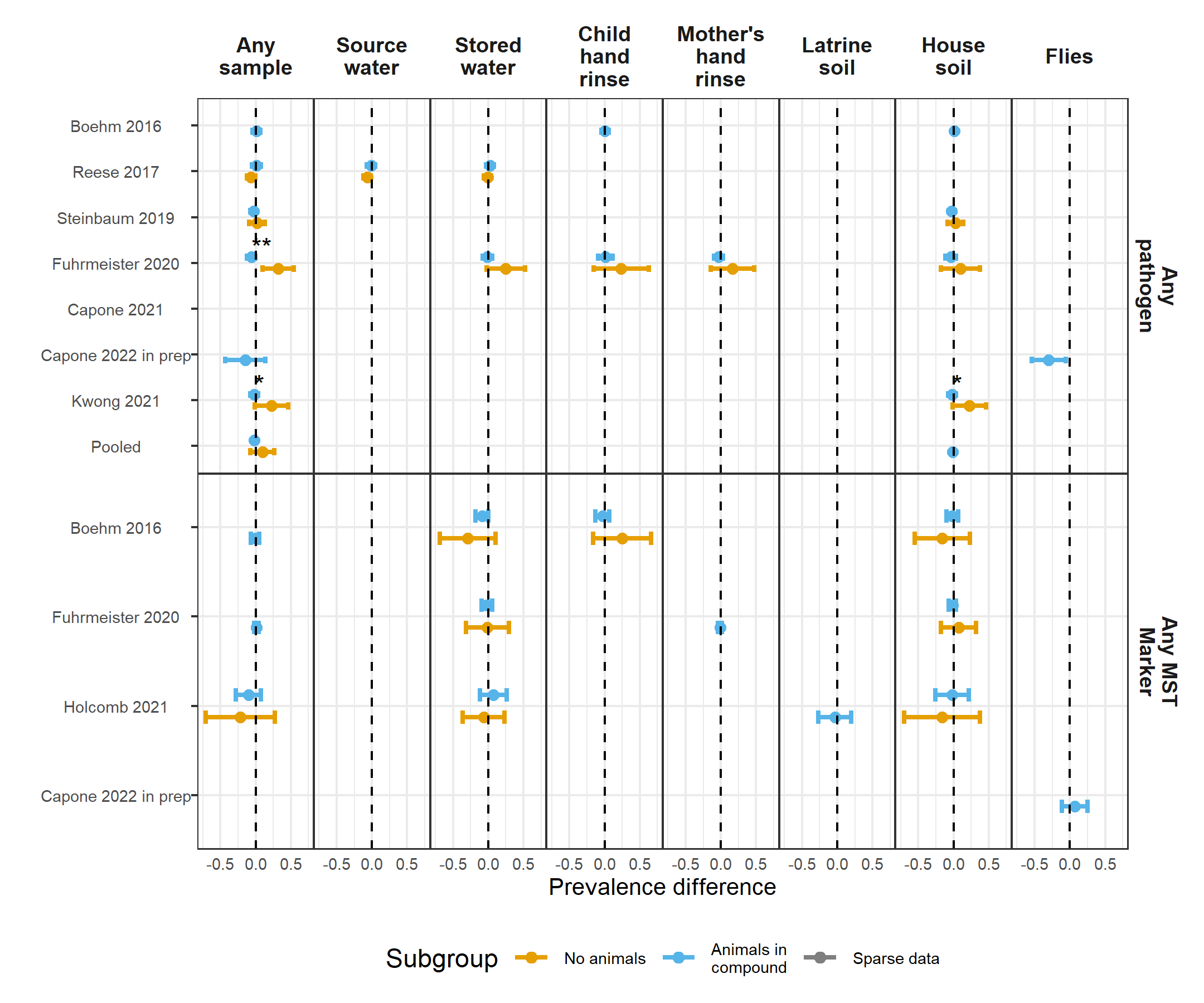


**Figure S4.** Distributions of abundance outcomes. The X-axes are displayed on the log-10 scale. Black vertical lines mark the means, and dashed lines mark the medians. Values below the limit of detection were imputed with with half the limit of detection and values below the limit of quantification were imputed with the midpoint between the limits of detections and quantification, leading to some bimodal distributions.

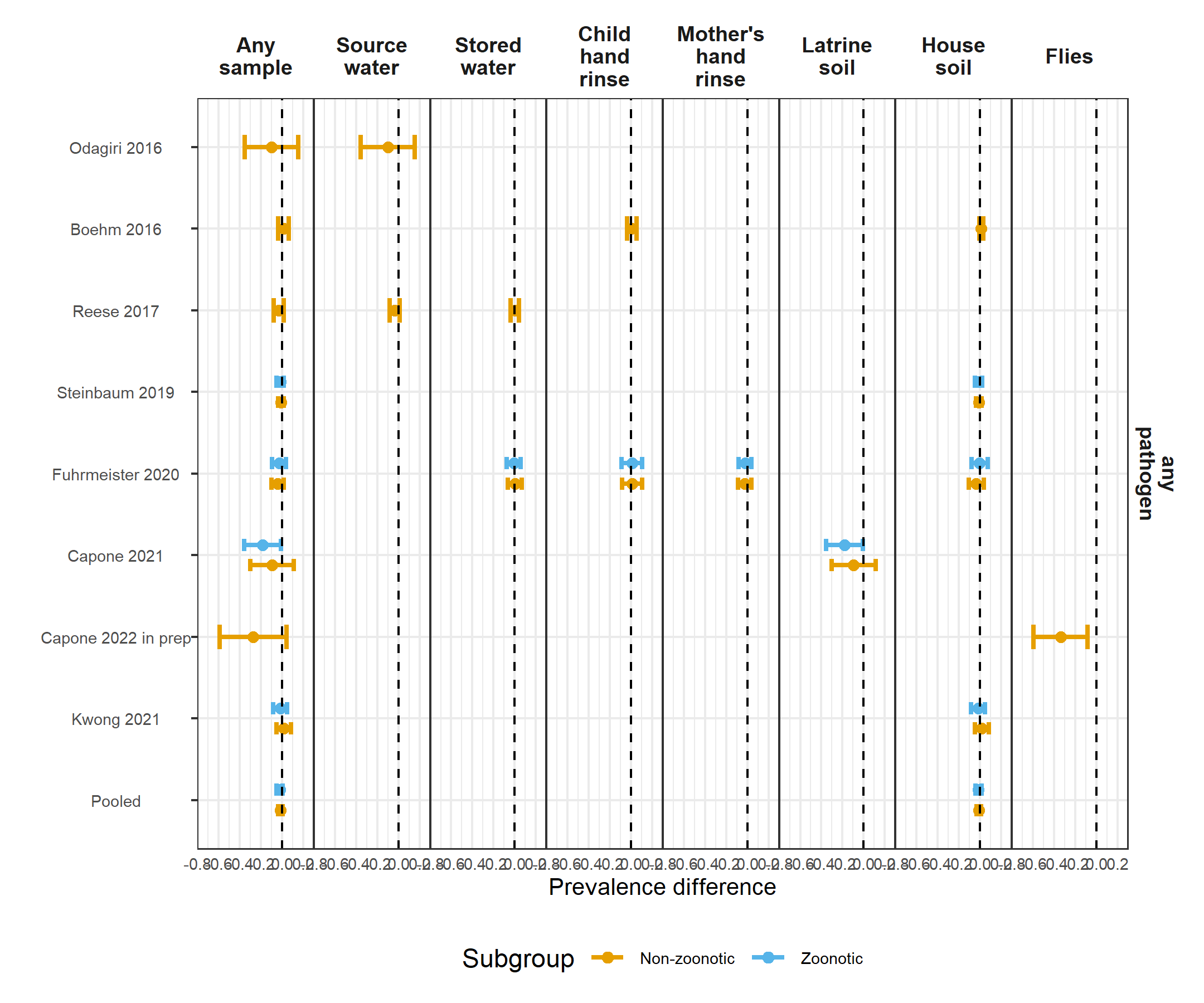


**Figure S5.**

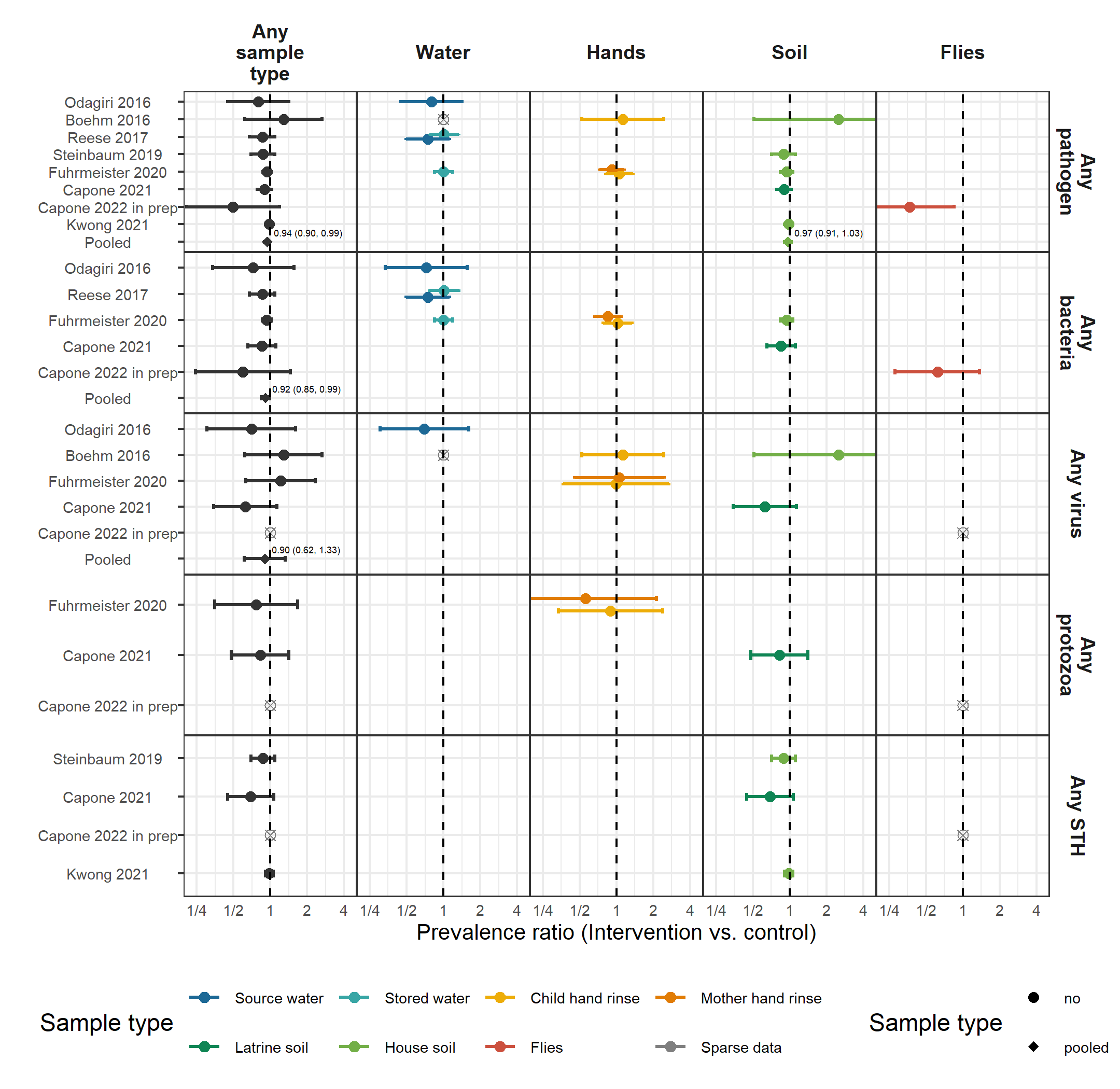
Forest plots of any enteropathogen prevalence differences or any MST prevalence differences between intervention and control arms, stratified by whether the sample was collected during the wet versus dry season (defined by the 6 months of highest average rainfall). Significant effect modification, as determined by the p-values on the regression model interaction term, is marked above points with asterisks (P < 0.05 = “\*”, P < 0.01 = “\*\*”, P < 0.001 = “\*\*\*”). Grey crossed points denote data that were too sparse to estimate a prevalence ratio (i.e., <10 positive observations).



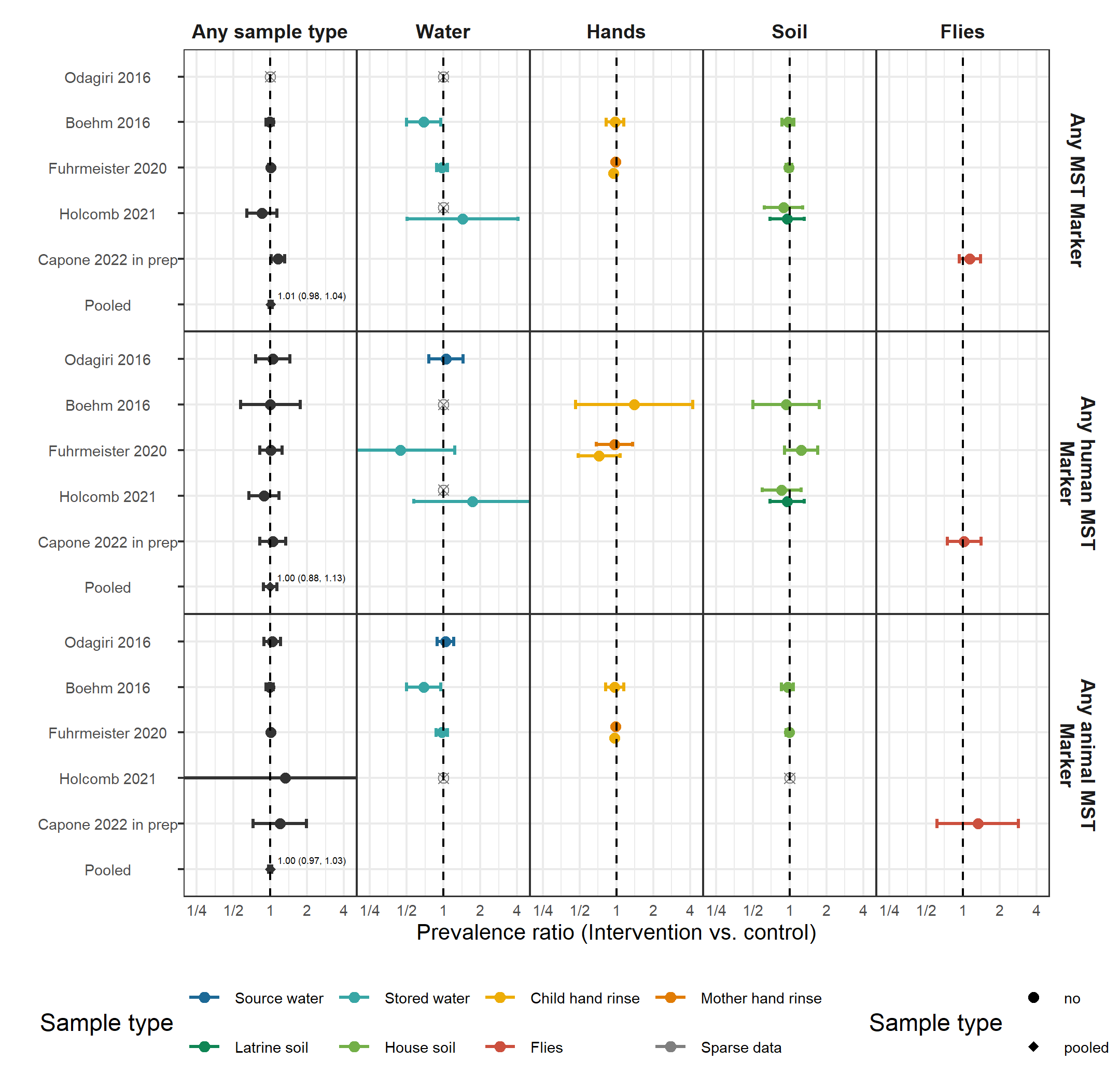
**Figure S6.** Forest plots of any enteropathogen prevalence differences or any MST prevalence differences between intervention and control arms, stratified by whether any animals were present in the compound. Significant effect modification, as determined by the p-values on the regression model interaction term, is marked above points with asterisks (P < 0.05 = “\*”, P < 0.01 = “\*\*”, P < 0.001 = “\*\*\*”). Grey crossed points denote data that were too sparse to estimate a prevalence ratio (i.e., <10 positive observations).



**Figure S7.** Forest plots of any enteropathogen prevalence differences or any MST prevalence differences between intervention and control arms, stratified by whether the pathogen is zoonotically transmitted. Grey crossed points denote data that were too sparse to estimate a prevalence ratio (i.e., <10 positive observations). Significant effect modification, as determined by the p-values on the regression model interaction term, is marked above points with asterisks (P < 0.05 = “\*”, P < 0.01 = “\*\*”, P < 0.001 = “\*\*\*”). Grey crossed points denote data that were too sparse to estimate a prevalence ratio (i.e., <10 positive observations).



**Figure S8.** Forest plots of unadjusted intervention effects on the prevalence of any enteropathogen or type of enteropathogen (any bacteria, any virus, any protozoa and any STH) in different types of environmental samples. Point estimates and confidence intervals are printed next to pooled estimates. Grey crossed points denote data that were too sparse to estimate a prevalence ratio (i.e., <10 positive observations).



**Figure S9.** Forest plots of unadjusted intervention effects on the prevalence of any MST marker or type of MST marker (human or animal MST markers) in different types of environmental samples. Point estimates and confidence intervals are printed next to pooled estimates. Grey crossed points denote data that were too sparse to estimate a prevalence ratio (i.e., <10 positive observations).