Acute illness among surfers following dry and wet weather seawater exposure

Benjamin F. Arnold, Kenneth C. Schiff, Ayse Ercumen, Jade Benjamin-Chung, Joshua A. Steele, John F. Griffith, Steven J. Steinberg, Paul Smith, Charles D. McGee, Richard Wilson, Chad Nelsen, Stephen B. Weisberg, John M. Colford, Jr.

American Journal of Epidemiology

Web Material

Web Table 1. STROBE Checklist

Strengthening the Reporting of Observational Studies in Epidemiology (STROBE)

Item	DESCRIPTION	REPORTED IN SECTION
Title and Abstract		
1a	Indicate the study's design with a commonly used term in the title or the abstract	Abstract (longitudinal cohort study)
1b	Provide in the abstract an informative and balanced summary of what was done and what was found	Abstract
Introduction		
Background/rationale		
2	Explain the scientific background and rationale for the investigation being reported	Introduction
Objectives		
3	State specific objectives, including any prespecified hypotheses	Introduction
Methods		
Study Design		
4	Present key elements of study design early in the paper	Methods (Study Design and Enrollment)
Setting		
5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	Methods (Setting, Study Design and Enrollment)
Participants		
6a	Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up	Methods (Study Design and Enrollment)
6b	For matched studies, give matching criteria and number of exposed and unexposed.	Not applicable
Variables		
7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	Methods (Outcome Definition and Measurement, Exposure Definition and Measurement, Statistical Analysis)
Data Sources and Management		
8	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	Methods (Outcome Definition and Measurement, Exposure Definition and Measurement)
Bias		

9	Describe any efforts to address potential sources of bias	Methods (Statistical Analysis)
Study Size		
10	Explain how the study size was arrived at	Web Appendix 2 (Power and sample size calculations)
Quantitative Variables		
11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	Methods (Statistical Analysis)
Statistical Methods		
12a	Describe all statistical methods, including those used to control for confounding	Methods (Statistical Analysis) Web Appendix 1 (Detailed stat. notation)
12b	Describe any methods used to examine subgroups and interactions	Methods (Statistical Analysis) Web Appendix 1 (Detailed stat. notation)
12c	Explain how missing data were addressed	Methods (Statistical Analysis)
12d	If applicable, explain how loss to follow-up was addressed	Methods (Statistical Analysis)
12e	Describe any sensitivity analyses	Methods (Statistical Analysis) Web Appendix 4 (Sensitivity analyses)
Results		
Participants		
13a	Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	Results (Study Population)
13b	Give reasons for non-participation at each stage	Refusal rates were not measurable
13c	Consider use of a flow diagram	Not applicable
Descriptive data		
14a	Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	Results (Study Population, Water Quality and Surfer Exposure) Table 1 Figs 1-2 Web Figs 1-2
14b	Indicate number of participants with missing data for each variable of interest	Table 1
14c	Summarise follow-up time (eg, average and total amount)	Results (Study Population, Water Quality and Surfer Exposure)
Outcome Data		

15	Report numbers of outcome events or summary measures over time	Table 2 Table 4 Web Table 2
Main Results		
16a	Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	Table 3 Table 4
16b	Report category boundaries when continuous variables were categorized	Not applicable
16c	If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	Web Fig 3 Web Fig 6
Other Analyses		
17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	Web Figs 3 - 7 Web Tables 2 - 5
Discussion		
Key Results		
18	Summarise key results with reference to study objectives	Discussion (Key Results)
Limitations		
19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	Discussion (Limitations)
Interpretation		
20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	Discussion (Interpretation)
Generalisability		
21	Discuss the generalisability (external validity) of the study results	Discussion (Interpretation)
Other Information		
Funding		
22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	Acknowledgements

Web Appendix 1. Statistical details

Analyses were pre-specified through the Open Science Framework (see repository for pre-registered analysis plan along with full replication files: https://osf.io/hvn7s). We defined incident episodes as the onset of symptoms preceded by ≥6 symptom-free days to increase the likelihood that separate episodes represented distinct infections (1,2). We treated an individual's first 6 days of follow-up time as "at risk" under the assumption that the individual did not have incident illness in the days immediately before the start of their recorded symptom history. We calculated incidence rates by dividing incident episodes by person-days in unexposed and exposed periods during follow-up. If participants missed weekly surveys during follow-up, we did not include those periods in the analysis. We measured the association between ocean exposure and subsequent illness using an incidence rate ratio (IRR). Let Y_{it} be a binary indicator equal to 1 if individual i is ill on day t (0 otherwise), let T_{it} be an indicator that participant i is at risk of illness on day t. Let E_{it} be a binary indicator of equal to 1 if individual i entered the ocean on day t (0 otherwise). Define $E^*_{it} = \max(E_{i,t-1}, \ldots, E_{i,t-3})$, which is a binary indicator of whether the individual entered the ocean in the three days prior to the outcome measurement on day t. Our first parameter of interest was the IRR associated with ocean exposure in the past three days ($E^* = 1$), averaged over potentially confounding covariates (X). We modeled illness for individual i on day t using the following log-linear rate model (3), subset to days at risk ($T_{it} = 1$):

$$\log \mathrm{E}[Y_{it} \mid E^*_{it}, X_{it}] = \alpha + \beta E^*_{it} + \gamma X_{it} \qquad (1)$$

where X_{it} is a vector of potential confounders included in adjusted analyses (details below). We estimated the IRR associated with ocean exposure from the model, $\exp(\beta)$, and used robust standard errors that accounted for repeated observations within individuals (4).

Our second research question examined whether ocean exposure increased illness rates more if exposure took place within three days of wet weather compared with exposure during dry weather. Let D_t be a count of days since it rained >0.25 cm in 24 hours, with $D_t = \{0, 1, 2, \ldots\}$. Let R_{it} be a binary indicator equal to 1 if individual i entered the ocean on day t and $D_t \leq 3$ (0 otherwise), indicating that a surf session took place within three days of rain. Define $R^*_{it} = \max(R_{i,t-1}, R_{i,t-2}, R_{i,t-3})$, a binary indicator of whether an individual had a wet weather exposure in the past three days. With E^*_{it} (an indicator of any ocean exposure in the past three days), we created a three level categorical exposure:

$$W_{it} = \begin{cases} E^*_{it} = 0, R^*_{it} = 0\colon \text{ unexposed} \\ E^*_{it} = 1, R^*_{it} = 0\colon \text{ dry} \\ E^*_{it} = 1, R^*_{it} = 1\colon \text{ wet} \end{cases}$$

We estimated a log-linear model, subset to days at risk ($T_{it} = 1$):

$$\log \mathrm{E}[Y_{it} \mid W_{it}, X_{it}] = \alpha + \beta_1 I(W_{it}=dry) + \beta_2 I(W_{it}=wet) + \gamma X_{it} \qquad (2)$$

where X_{it} are covariates in adjusted models. We estimated separate IRRs from the model for surf exposure during dry versus unexposed periods, $\exp(\beta_1)$, for surf exposure during wet versus unexposed periods, $\exp(\beta_2)$, and for wet versus dry periods, $\exp(\beta_2 - \beta_1)$. For each outcome, we calculated a test of trend in the IRRs for dry and wet weather exposures (not pre-specified), in which the test for log-linear trend in incidence rates was significant if the coefficient β_2 differed from zero (5).

We estimated the association between fecal indicator bacteria levels and illness using the subset of surf sessions matched to water quality indicator measurements at the sentinel beaches using $\log_{10}$ continuous indicator levels, F_{it} . For surfers with a single day of exposure matched to indicator levels in the past three days, F_{it} equaled the daily geometric mean value on the exposed day. For surfers with multiple exposures matched to indicator levels in the past three days, we calculated the mean concentration weighted by the number of hours spent in the water on each day. We modeled the relationship between indicator levels and illness for individual *i* on day *t* using a log-linear model, subset to days at risk ($T_{it} = 1$):

$$\log \mathrm{E}[Y_{it} \mid F_{it}, X_{it}] = \alpha + \delta F_{it} + \gamma \mathbf{X}_{it} \tag{3}$$

where $\exp(\delta)$ estimates the IRR associated with a 1-$\log_{10}$ increase in indicator level. We also estimated the IRR associated with values above versus below USEPA *Enterococcus* regulatory guidelines (6) by replacing F_{it} in equations 3 and 4 with an indicator equal to 1 if F_{it} exceeded 35 CFU/100ml or, in a second definition, if any single sample on the exposure day exceeded 104 CFU/100ml. In the water quality analysis, we also hypothesized that the relationship between fecal indicator bacteria and illness could be modified by whether it was dry or wet weather exposure. We allowed the exposure-response relationship to vary by exposure during dry and wet weather by including an indicator for wet weather exposure in the past three days, R^*_{it} , and an interaction term in the model:

$$\log \mathrm{E}[Y_{it} \mid F_{it}, W^*_{it}, X_{it}] = \alpha + \delta_1 F_{it} + \delta_2 R^*_{it} + \delta_3 F_{it} R^*_{it} + \gamma \mathbf{X}_{it} \tag{4}$$

<u>Potential confounders</u>: We selected potential confounders that could be either a cause of seawater exposure, a cause of illness, or both (7). We controlled for the following time-invariant

potential confounders measured at enrollment: age, sex, education, employment status, household income, years the individual had surfed, reported behavior of typically avoiding the ocean following wet weather, surfboard length, mode of enrollment (beach vs. web), and chronic health conditions included only for the corresponding outcomes: ear problems, sinus problems, gastrointestinal conditions, respiratory conditions, skin conditions. We also controlled for time-varying potential confounders: entered the ocean for an activity other than surfing, any illness symptoms in the week preceding the risk window, day of recall, day of the week, and rainfall total during the past three days. In the adjusted overall seawater exposure analysis and water quality analysis we considered an indicator of wet weather in the past three days; in the water quality analysis we also considered an indicator for sentinel beach and an indicator for whether the individual surfed at beaches other than our two sentinel beaches in the same three-day period as their sentinel beach exposure. From this set of potential confounders, we retained those that had a univariate association with the outcome, defined as a likelihood ratio test *P*-value <0.20 in an unadjusted model. For categorical variables, we included a “missing” category for missing values.

Web Appendix 2. Power and sample size calculations

The sample size for the study was developed in two stages because little was known about outcome or exposure prevalence in the surfer population. In year 1, we aimed to enroll 100-200 surfers and follow them for up to 12 weeks to collect exposure and illness information, as well as fecal indicator bacteria levels. Using exposure and outcome information from the initial, smaller cohort in year 1, we then calculated sample size and power for the full study and this informed enrollment targets for year 2.

During the first year of the study, we enrolled 162 individuals and observed 12 incident cases of gastrointestinal illness from 2,310 days at risk -- an incidence rate of 5 episodes per 1,000 person-days. We used a standard sample size equation for the comparison of two incidence rates (8): $y = (z_{\alpha/2} + z_{\beta})^2 (\lambda_0 + \lambda_1) / (\lambda_0 - \lambda_1)^2$, where y is the number of person-days required in each exposure category, $z_{\alpha/2}$ and z_{β} are standard normal distribution values corresponding to upper tail probability values $\alpha/2$ and β (we set α=0.05 and β=0.2), and λ_0 and λ_1 are incidence rates in the unexposed and exposed periods. Assuming a rate of 5 episodes per 1,000 person-days during unexposed periods (λ_0=0.005), the Table below summarizes the number of person-days of observation in each exposure group required to detect different magnitudes of effect, as measured by the incidence rate ratio (IRR).

In year 1 of the study, 55% of the days of observation were exposed because surfers entered the ocean frequently. However, only 13% of the days of observation were classified as wet weather exposure because it was a drought year.

Given this, we expected that a total of 2,408 / 0.13 = 18,520 person-days of observation would be sufficient to detect an IRR of 1.50 or greater in wet weather exposed versus unexposed periods.

IRR	Person-days of observation in each exposure group (*y*)	Total person-days required, assuming 13% of days are wet weather exposure (*y* / 0.13)
1.2	13,242	101,859
1.3	6,153	47,328
1.4	3,611	27,780
1.5	2,408	18,520

For associations between $\log_{10}$ *Enterococcus* and incident illness we used a simulation-based approach (9). The simulation resampled the empirical distribution of water quality measurements from year 1 for 200 surfers with different lengths of follow-up and calculated a predicted probability of incident gastrointestinal illness on each day using the rate in the unexposed periods from year 1 and an increased rate following exposure that corresponded to different effect sizes. For a given strength of association (IRR), we then increased the length of follow-up until >80% of the 1,000 simulations had a P<0.05 (equivalent to α=0.05 and β=0.2). Simulations showed that 3,000 days of follow-up matched to water quality indicator measurements at sentinel beaches would provide >80% power to estimate an IRR of 1.75 or greater for a $\log_{10}$ increase in *Enterococcus* levels.

Web Appendix 3. Conversion of incidence rates into cumulative incidence

The longitudinal design with varying lengths of follow-up and varying exposure periods meant that the natural measure of illness was incidence rates (episodes / person-days) (3). However, federal water quality guidelines and quantitative microbial risk assessment models measure illness in units of cumulative incidence or “risk” (episodes / person) for gastrointestinal illness (6). We converted marginally adjusted incidence rate estimates from log-linear models described above into 3-day cumulative incidence using the density method (10). We compared cumulative incidence during dry and wet weather exposure periods to unexposed periods using the difference in cumulative incidence (“risk difference” [RD]), and estimated standard errors and 95% confidence intervals for the RD using the delta method (11). We used a 3-day cumulative incidence because incidence rates were measured over 3-day periods following exposure -- the high frequency of exposure made longer follow-up periods infeasible. In California swimmer cohorts, the majority of excess cases of gastrointestinal illness occurred in the 1-2 days following ocean exposure; for this reason, a 3-day RD should be a reasonable approximation of the RD calculated over a longer 10-12 day period, as measured in past swimmer cohort studies (12–14).

Web Appendix 4. Sensitivity analyses

The primary analysis defined wet weather exposure as periods within 0-3 days following rainfall, consistent with current beach posting guidelines in California that warn recreators to avoid water contact for 72 hours after rainfall. In a sensitivity analysis we changed the length of the wet weather window in daily increments from 0 to 5 days following rainfall to determine if shorter windows were associated with larger increase in illness rates.

The technical report associated with this study (https://osf.io/hvn7s) includes additional sensitivity analyses. First, we stratified incidence rates by storm size. We found evidence for higher incidence rates associated with ocean exposure following larger storms compared with smaller storms, but the analysis relied upon relatively small sample sizes with the higher level of wet weather stratification and therefore should be viewed as supportive but exploratory. Second, we examined the effect of excluding individuals from the analysis who submitted >1 survey per day (to reduce potential measurement bias) or who did not report the precise location of their ocean exposure (to reduce potential exposure misclassification). Excluding these subgroups of the population from the analysis did not change our inference.

Web Appendix 5. Negative control exposure analyses

We matched fecal indicator bacteria levels (*Enterococcus*, fecal coliforms, total coliforms) measured at one of the sentinel beaches to illness measurements by date, randomly assigning either the Ocean Beach value or the Tourmaline Surfing Park value to each matched observation. We then excluded from the dataset observations that were within five days of any seawater exposure. The negative control exposure analysis was designed to ensure that the person-time included could not plausibly be influenced by pathogens in seawater, but could be influenced by unmeasured sources of confounding -- such as seasonal epidemics of enteric pathogens like norovirus.

We estimated the association between fecal indicator bacteria levels and incident illness, allowing the association to vary by dry and wet weather exposure by including an interaction term between the indicator levels and a wet weather period indicator, as in the main analysis. In this negative control exposure analysis, there should be no plausible relationship between fecal indicator bacteria levels and subsequent illness unless there is bias from unobserved confounding or measurement error (15,16).

Web Appendix 6. Comparison of gastrointestinal illness rates with summer cohorts

To help contextualize the gastrointestinal illness rates measured in this study we completed a supplemental analysis to more directly compare illness rates with those measured in summer swimmer cohorts. The present study used a different design that past recreational swimmer cohorts conducted in California (12–14,17,18). Past swimmer cohorts enrolled beachgoers and then measured cumulative incident illness over 10-12 days after their single exposure. Such measurement was infeasible among surfers because of their frequent exposure (median of 2 times per week). For this reason, the present study estimated daily incidence rates during unexposed and exposed periods of follow-up -- the most natural measure of disease given the design. This difference in design and measure of illness complicates direct comparisons of illness between this study and past swimmer cohorts. A second difference in design that complicates direct comparison is that the present study limited enrollment to adults, whereas past swimmer cohorts included many children who have higher rates of gastrointestinal illness (18).

We had access to participant data from four California swimmer cohorts (18), and this enabled us to derive estimates of gastrointestinal illness rates from the past studies that were more comparable to those estimated in the surfer cohort. We subset the four California cohorts (Avalon, Doheny, Malibu, Mission Bay) to adults (18 years or older) and calculated incidence rates over the first 3 days of follow-up -- a period after exposure comparable to the present study. We calculated incidence rates separately for non-swimmers (individuals with no water contact) and swimmers with head immersion exposure. We compared these gastrointestinal incidence rates with rates among surfers during unexposed and exposed periods in the present study.

Web Appendix References

1. Colford JM, Wade TJ, Sandhu SK, et al. A randomized, controlled trial of in-home drinking water intervention to reduce gastrointestinal illness. *Am. J. Epidemiol.* 2005;161(5):472–482.

2. Colford JM, Hilton JF, Wright CC, et al. The Sonoma Water Evaluation Trial: A Randomized Drinking Water Intervention Trial to Reduce Gastrointestinal Illness in Older Adults. *Am. J. Public Health*. 2009;99(11):1988–1995.

3. Rothman KJ, Sander Greenland, Lash TL. Modern Epidemiology. 3rd ed. Philadelphia: Lippincott Williams and Wilkins; 2008.

4. Huber PJ. The behavior of maximum likelihood estimates under nonstandard conditions. In: *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*. books.google.com; 1967:221–233.

5. Vittinghoff E, Glidden DV, Shiboski SC, et al. Regression methods in biostatistics: linear, logistic, survival, and repeated measures models. 2nd ed. Springer Science & Business Media; 2012.

6. USEPA. Recreational Water Quality Criteria. United States Environmental Protection Agency Office of Water; 2012.

7. VanderWeele TJ, Shpitser I. A new criterion for confounder selection. *Biometrics*. 2011;67(4):1406–1413.

8. Hayes RJ, Bennett S. Simple sample size calculation for cluster-randomized trials. *Int. J. Epidemiol.* 1999;28(2):319–326.

9. Arnold B, Hogan D, Colford J, et al. Simulation methods to estimate design power: an overview for applied research. *BMC Med. Res. Methodol.* 2011;11(1):94.

10. Kleinbaum DG, Kupper LL, Morgenstern H. Epidemiologic Research: Principles and Quantitative Methods. New York: John Wiley & Sons, Inc.; 1982.

11. Wasserman L. All of statistics: a concise course in statistical inference. Springer; 2004.

12. Colford JM, Schiff KC, Griffith JF, et al. Using rapid indicators for Enterococcus to assess the risk of illness after exposure to urban runoff contaminated marine water. *Water Res.* 2012;46(7):2176–2186.

13. Arnold BF, Schiff KC, Griffith JF, et al. Swimmer illness associated with marine water exposure and water quality indicators: impact of widely used assumptions. *Epidemiology*. 2013;24(6):845–853.

14. Yau VM, Schiff KC, Arnold BF, et al. Effect of submarine groundwater discharge on bacterial indicators and swimmer health at Avalon Beach, CA, USA. *Water Res.* 2014;59:23–36.

15. Lipsitch M, Tchetgen ET, Cohen T. Negative controls: a tool for detecting confounding and bias in observational studies. *Epidemiology*. 2010;21(3):383–388.

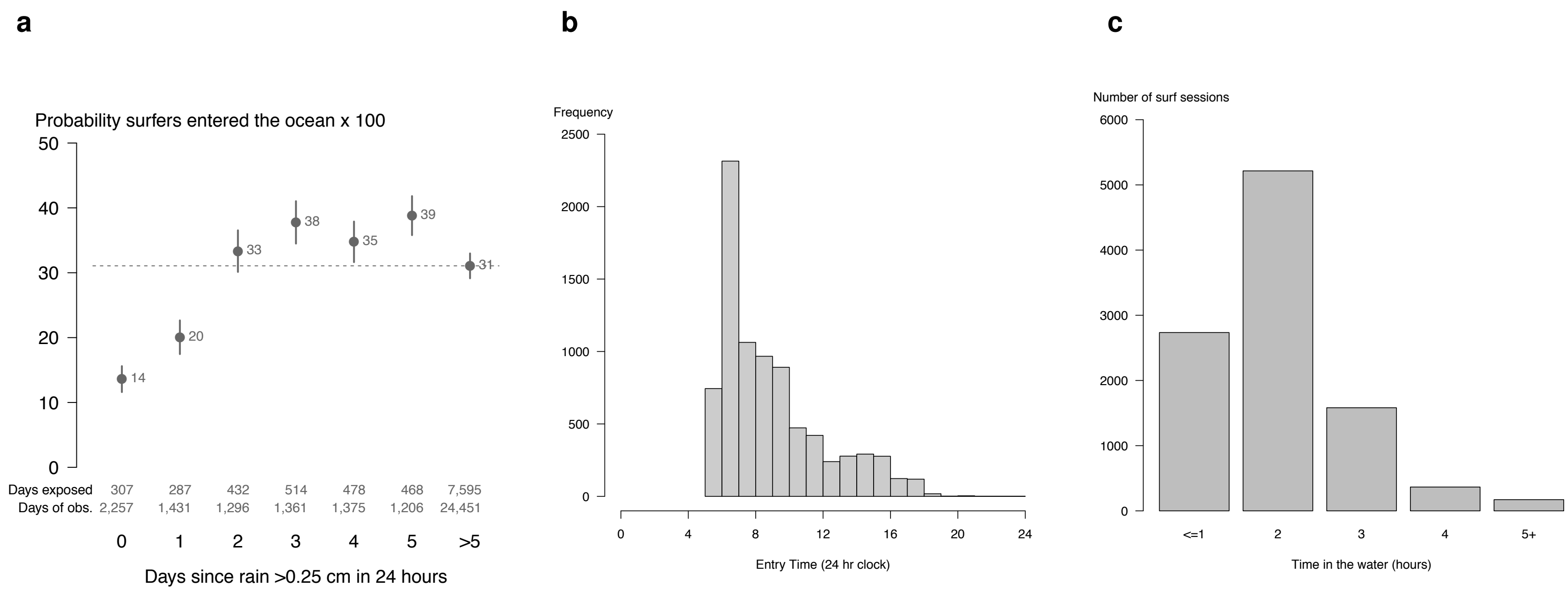
16. Arnold BF, Ercumen A, Benjamin-Chung J, et al. Negative controls to detect selection bias and

measurement bias in epidemiologic studies. *Epidemiology*. 2016;27(5):637–641.

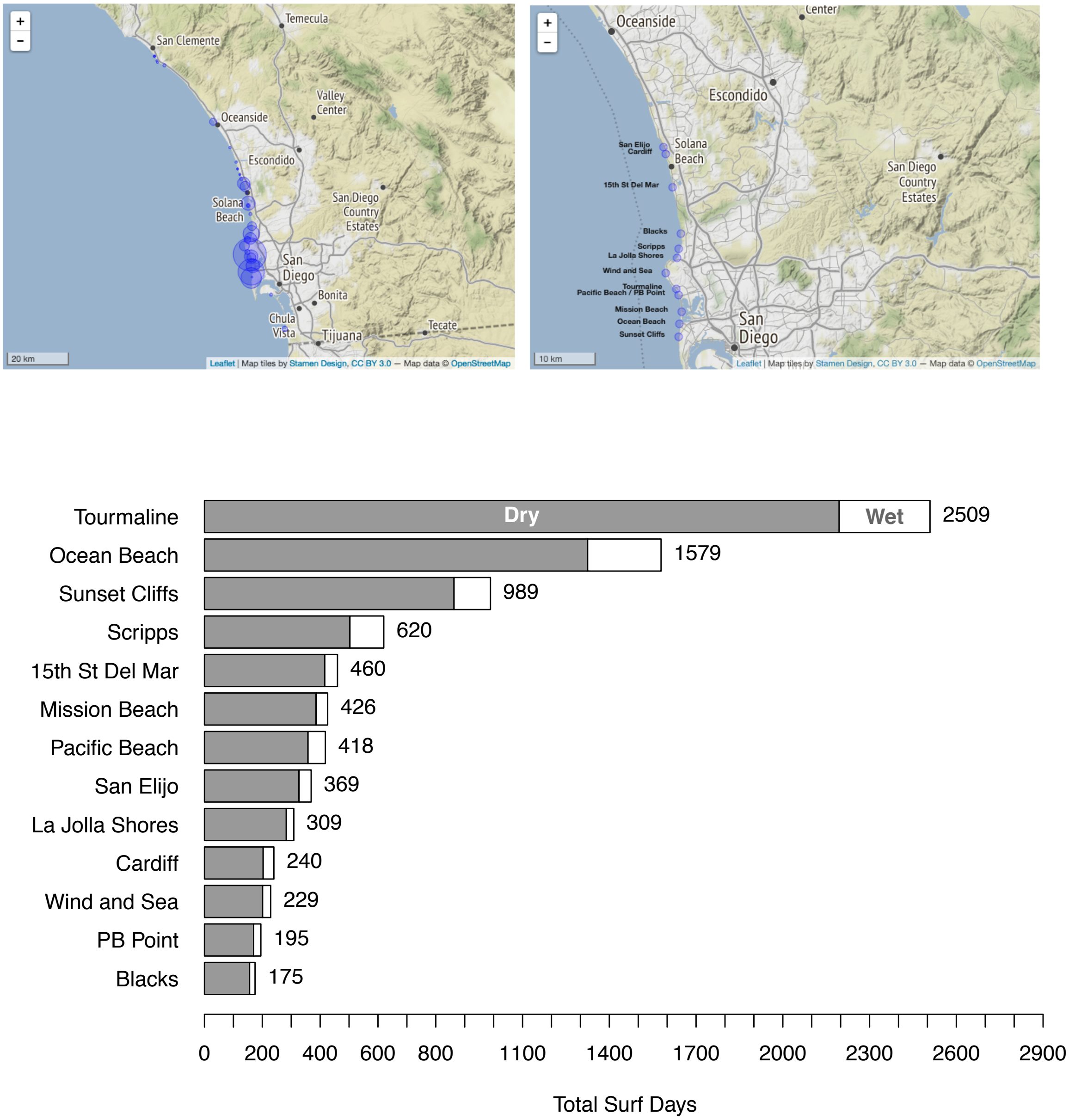
17. Colford JM, Wade TJ, Schiff KC, et al. Water quality indicators and the risk of illness at beaches with nonpoint sources of fecal contamination. *Epidemiology*. 2007;18(1):27–35.

18. Arnold BF, Wade TJ, Benjamin-Chung J, et al. Acute Gastroenteritis and Recreational Water: Highest Burden Among Young US Children. *Am. J. Public Health*. 2016;106(9):1690–1697.

Additional Web Tables and Figures



Web Figure 1: Surfer exposure during follow-up, summarized from 654 surfers (10,081 surf sessions) in the San Diego, CA region during the winters of 2013-14 and 2014-15. **a**, Probability that surfers entered the ocean, stratified by days since precipitation >0.25 cm in 24 hours. Vertical lines indicate robust 95% confidence intervals and the dashed line marks the probability for >5 days after rain. **b**, Distribution of ocean entry times. **c**, Distribution of time spent in the ocean, rounded to hours.



Web Figure 2: Distribution of surf days for surf locations in San Deigo county (top left panel), where the area of each circle is scaled by the total surf days. The 13 most common surf locations (top right panel) represented 85% (8,518/10,081) of surf days observed in the study. The bottom panel shows the distribution of surf days at the 13 most popular locations. Wet weather was defined as >0.25 cm of rain in 24 hours. An interactive version of the maps is located in the study's open science framework repository: `https://osf.io/bfxq4`

Web Table 2: Incident illness and incidence rate ratios (IRR) associated with ocean exposure in the past three days among surfers in San Diego, CA (2013-14 and 2014-15 winters).

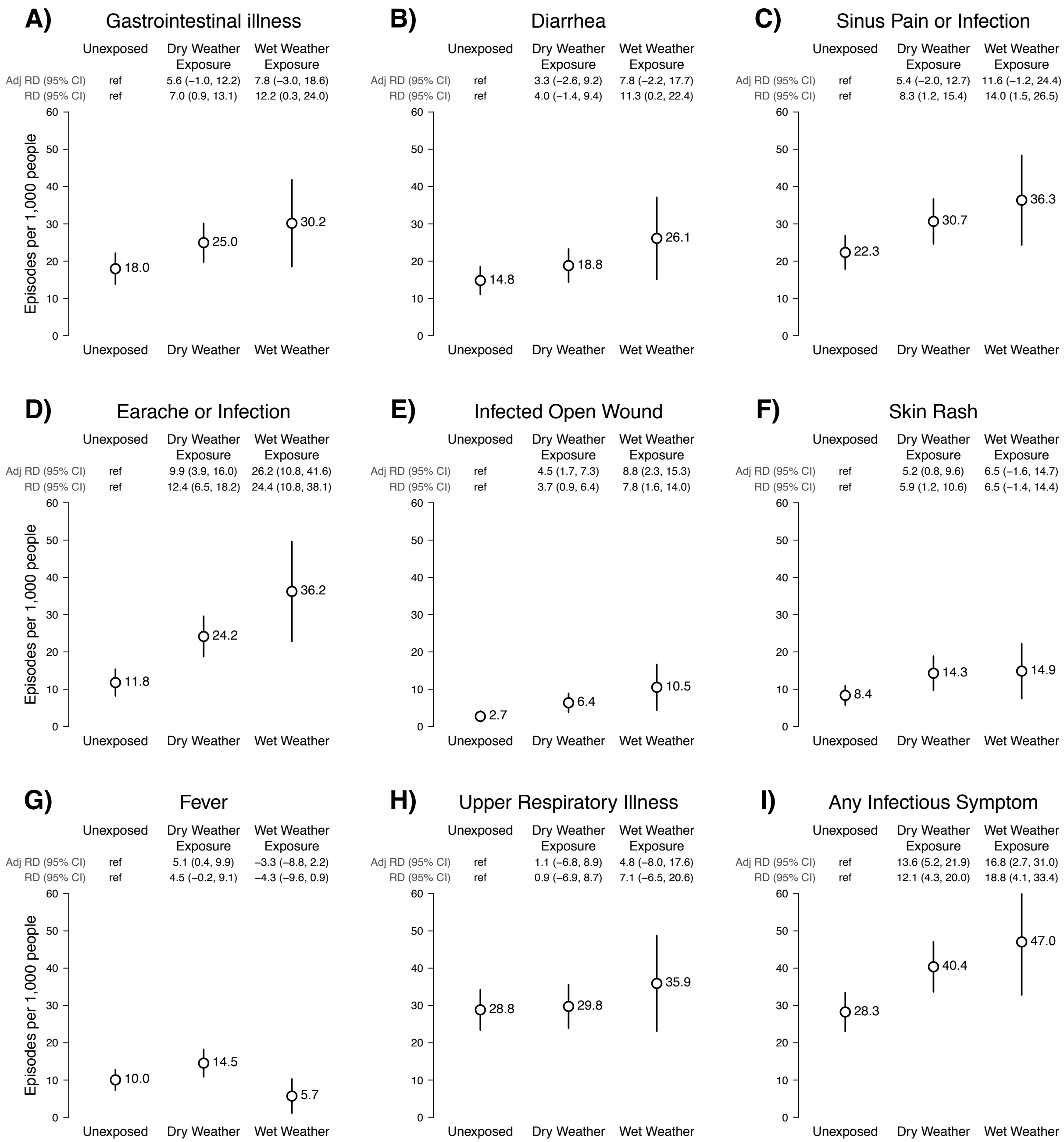
	Unexposed Periods		Ocean Exposure, Past 3 Days		Unadjusted	Adjusted [b]
	Episodes/days at risk	Rate [a]	Episodes/days at risk	Rate [a]	IRR (95% CI)	IRR (95% CI)
Gastrointestinal illness	90/14,884	6.0	147/16,806	8.7	1.45 (1.10, 1.91)	1.33 (0.99, 1.78)
Diarrhea	75/15,086	5.0	115/16,970	6.8	1.36 (1.00, 1.86)	1.29 (0.93, 1.79)
Sinus pain or infection	109/14,475	7.5	176/16,389	10.7	1.43 (1.11, 1.83)	1.28 (0.98, 1.68)
Earache or infection	59/14,931	4.0	148/16,626	8.9	2.25 (1.60, 3.16)	2.18 (1.48, 3.20)
Infection of open wound	14/15,456	0.9	41/17,199	2.4	2.63 (1.45, 4.77)	3.28 (1.68, 6.43)
Skin rash	42/15,024	2.8	81/16,757	4.8	1.73 (1.19, 2.51)	1.67 (1.15, 2.42)
Fever	51/15,156	3.4	75/17,290	4.3	1.29 (0.89, 1.87)	1.39 (0.93, 2.07)
Upper respiratory illness [c]	117/12,001	9.7	142/13,568	10.5	1.07 (0.83, 1.39)	1.07 (0.83, 1.37)
Any infectious symptom [d]	138/14,445	9.6	228/16,102	14.2	1.48 (1.19, 1.85)	1.52 (1.20, 1.93)

a Episodes per 1,000 person-days

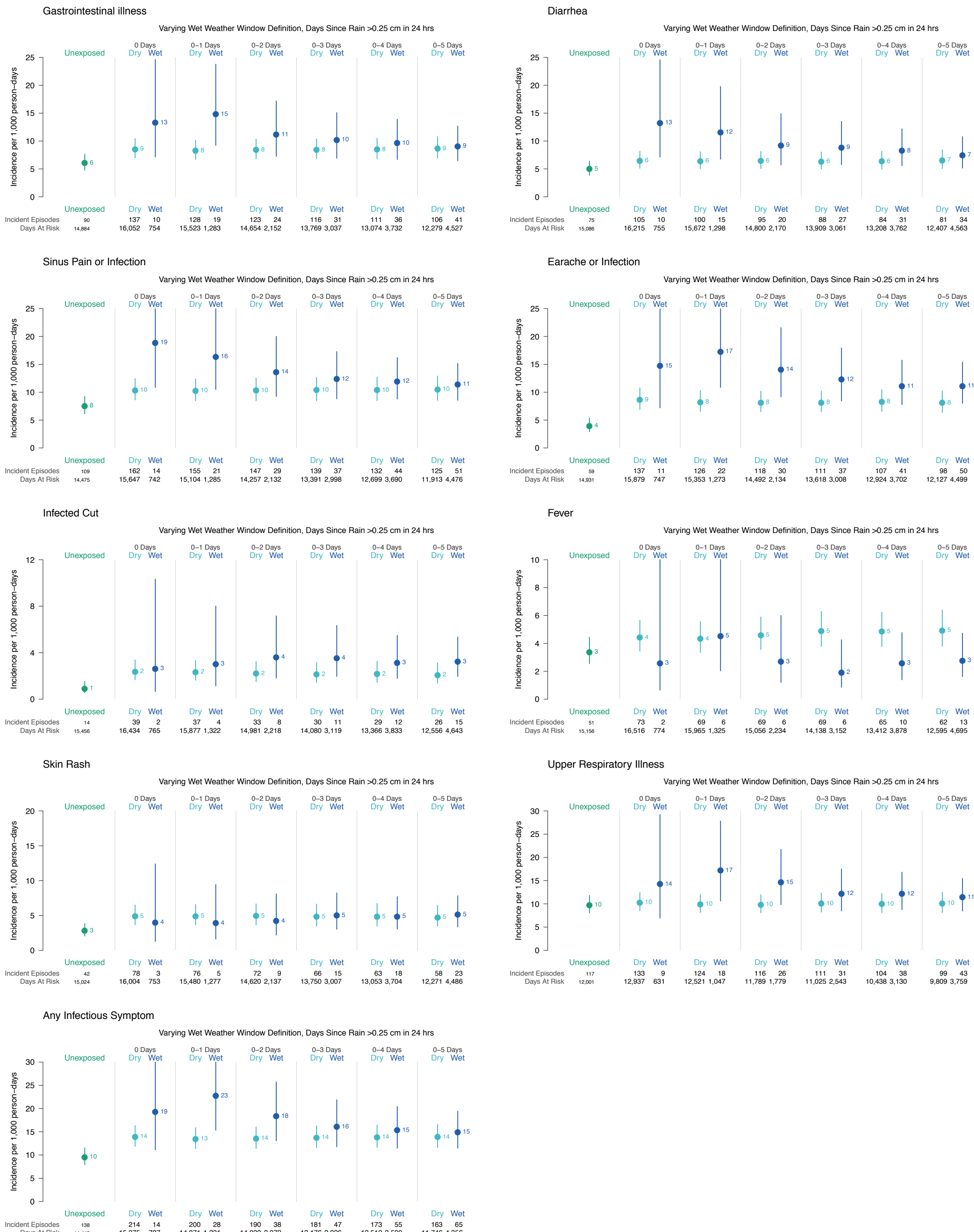
b Adjusted for a range of covariates (see statsistical methods for details)

c Only measured in year 2

d Includes gastrointestinal illness, eye infections, infected wounds, and fever.



Web Figure 3: Three-day cumulative incidence of illness among surfers associated with dry and wet weather exposure in San Diego, CA during the winters of 2013-14 and 2014-15. Unadjusted and adjusted risk differences (RD) compare cumulative incidence in the three days following ocean exposure during dry or wet weather with three-day cumulative incidence during unexposed periods. Wet weather was defined as >0.25 cm of rain in 24 hours.



Web Figure 4: Sensitivity analysis of wet weather exposure period definition on illness incidence rates among surfers in San Diego, CA (2013-14 and 2014-15 winters). Wet weather was defined as >0.25 cm of rain in 24 hours. Incidence rates for dry and wet weather were re-calculated for varying lengths of wet weather window. The primary analysis used a period of 0-3 days.

Web Table 3: Incidence rate ratios (IRR) associated with fecal indicator bacteria among surfers exposed at Tourmaline Surfing Park and Ocean Beach in San Diego, CA (2013-14 and 2014-15 winters). The IRR is associated with either a log10 increase in indicator bacteria (*Enterococcus* , fecal coliforms, total coliforms), or exposure above versus below *Enterococcus* regulatory guidelines of 35 or 104 colony forming units (CFU) per 100/ml

		Episodes/ days at risk	Unadjusted IRR	(95% CI)	Adjusted [e] IRR	(95% CI)
***Enterococcus* $\log_{10}$**	Gastrointestinal illness	40 / 5548	1.23	(0.74 , 2.03)	1.04	(0.63 , 1.72)
	Diarrhea	33 / 5590	1.54	(0.94 , 2.51)	1.36	(0.83 , 2.21)
	Sinus pain or infection	63 / 5392	1.61	(1.13 , 2.28)	1.27	(0.85 , 1.90)
	Earache or infection	52 / 5507	0.94	(0.54 , 1.62)	0.91	(0.53 , 1.56)
	Infection of open wound	15 / 5692	2.67	(1.39 , 5.13)	3.24	(1.66 , 6.31)
	Skin rash	24 / 5497	1.23	(0.65 , 2.31)	0.88	(0.39 , 1.98)
	Fever	24 / 5708	1.28	(0.74 , 2.22)	1.29	(0.74 , 2.24)
	Upper respiratory illness [c]	52 / 4769	1.28	(0.80 , 2.03)	1.12	(0.71 , 1.78)
	Any infectious symptom [d]	67 / 5344	1.51	(1.04 , 2.19)	1.42	(0.95 , 2.11)
***Enterococcus* > 35 CFU [a]**	Gastrointestinal illness	40 / 5548	1.51	(0.76 , 3.02)	1.20	(0.59 , 2.44)
	Diarrhea	33 / 5590	2.08	(1.01 , 4.32)	1.77	(0.86 , 3.65)
	Sinus pain or infection	63 / 5392	1.80	(1.05 , 3.10)	1.35	(0.72 , 2.55)
	Earache or infection	52 / 5507	1.32	(0.71 , 2.47)	1.31	(0.72 , 2.37)
	Infection of open wound	15 / 5692	1.46	(0.52 , 4.09)	1.81	(0.64 , 5.15)
	Skin rash	24 / 5497	1.76	(0.74 , 4.16)	1.22	(0.43 , 3.47)
	Fever	24 / 5708	1.44	(0.58 , 3.55)	1.49	(0.62 , 3.62)
	Upper respiratory illness [c]	52 / 4769	0.97	(0.52 , 1.80)	0.80	(0.43 , 1.47)
	Any infectious symptom [d]	67 / 5344	1.47	(0.89 , 2.45)	1.35	(0.76 , 2.39)
***Enterococcus* > 104 CFU [b]**	Gastrointestinal illness	40 / 5548	1.27	(0.66 , 2.44)	1.05	(0.53 , 2.06)
	Diarrhea	33 / 5590	1.75	(0.87 , 3.49)	1.53	(0.75 , 3.13)
	Sinus pain or infection	63 / 5392	1.85	(1.09 , 3.14)	1.39	(0.75 , 2.57)
	Earache or infection	52 / 5507	1.28	(0.72 , 2.28)	1.25	(0.71 , 2.20)
	Infection of open wound	15 / 5692	2.76	(1.03 , 7.44)	3.28	(1.16 , 9.23)
	Skin rash	24 / 5497	1.21	(0.51 , 2.85)	0.74	(0.27 , 2.07)
	Fever	24 / 5708	1.42	(0.58 , 3.49)	1.45	(0.59 , 3.55)
	Upper respiratory illness [c]	52 / 4769	1.09	(0.60 , 1.98)	0.89	(0.50 , 1.59)
	Any infectious symptom [d]	67 / 5344	1.79	(1.10 , 2.92)	1.69	(1.01 , 2.84)
Fecal coliform $\log_{10}$	Gastrointestinal illness	40 / 5548	1.40	(0.79 , 2.46)	1.14	(0.64 , 2.05)
	Diarrhea	33 / 5590	1.76	(1.00 , 3.07)	1.56	(0.89 , 2.73)
	Sinus pain or infection	63 / 5392	1.85	(1.22 , 2.81)	1.31	(0.83 , 2.08)
	Earache or infection	52 / 5507	1.09	(0.61 , 1.95)	1.02	(0.57 , 1.84)
	Infection of open wound	15 / 5692	2.99	(1.47 , 6.07)	4.16	(1.84 , 9.41)
	Skin rash	24 / 5497	1.39	(0.69 , 2.81)	0.91	(0.36 , 2.28)
	Fever	24 / 5708	1.21	(0.57 , 2.60)	1.27	(0.61 , 2.65)
	Upper respiratory illness [c]	52 / 4769	1.46	(0.79 , 2.70)	1.18	(0.66 , 2.13)
	Any infectious symptom [d]	67 / 5344	1.73	(1.15 , 2.60)	1.66	(1.07 , 2.57)
Total coliform $\log_{10}$	Gastrointestinal illness	40 / 5548	1.17	(0.68 , 2.00)	1.03	(0.63 , 1.66)
	Diarrhea	33 / 5590	1.15	(0.61 , 2.17)	1.04	(0.58 , 1.88)
	Sinus pain or infection	63 / 5392	1.73	(1.17 , 2.57)	1.50	(0.95 , 2.35)
	Earache or infection	52 / 5507	1.29	(0.75 , 2.21)	1.26	(0.72 , 2.21)
	Infection of open wound	15 / 5692	2.78	(1.35 , 5.73)	3.46	(1.78 , 6.75)
	Skin rash	24 / 5497	1.29	(0.72 , 2.29)	1.07	(0.51 , 2.27)
	Fever	24 / 5708	1.38	(0.75 , 2.52)	1.47	(0.81 , 2.67)
	Upper respiratory illness [c]	52 / 4769	1.31	(0.80 , 2.16)	1.20	(0.74 , 1.93)
	Any infectious symptom [d]	67 / 5344	1.70	(1.14 , 2.52)	1.65	(1.10 , 2.49)

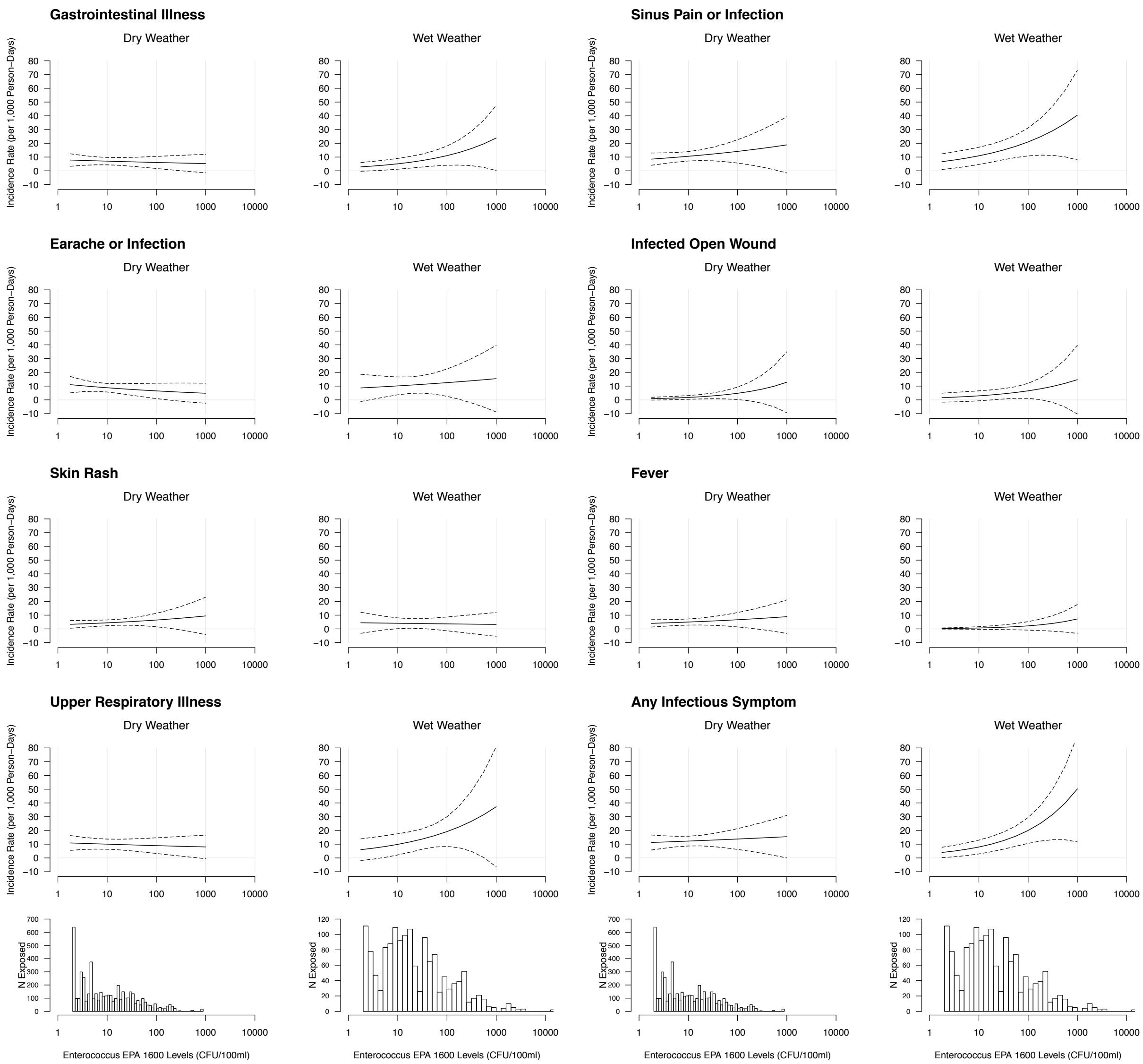
[a] Daily geomean > 35 CFU at any point during three-day follow-up

[b] Any single sample > 104 CFU at any point during three-day follow-up

[c] Only measured in year 2

[d] Includes gastrointestinal illness, diarrhea, vomiting, eye infections, infected cuts and fever

[e] Adjusted for a range of covariates (see statistical methods for details)



Web Figure 5: Incidence rates associated with *Enterococcus* levels measured during dry and wet weather periods, predicted from a log-linear model among surfers at Tourmaline Surfing Park and Ocean Beach, San Diego, CA (2013-14 and 2014-15 winters). Wet weather was defined as >0.25 cm of rain in 24 hours. Dashed lines indicate model-based 95% confidence intervals and histograms show the distribution of *Enterococcus* exposure in the population during dry and wet weather periods.

Web Table 4: Incidence rate ratios (IRR) associated with exposure above versus below *Enterococcus* regulatory guideline values, stratified by exposure during dry and wet weather, among surfers exposed at Tourmaline Surfing Park and Ocean Beach in San Diego, CA (2013-14 and 2014-15 winters).

		Dry Weather					**Wet Weather**						
		Episodes/	**Unadjusted**		**Adjusted** [e]		**Episodes/**	**Unadjusted**			**Adjusted** [e]		
		days at risk	**IRR**	**(95% CI)**	**IRR**	**(95% CI)**	**days at risk**	**IRR**	**(95% CI)**	**p** [f]	**IRR**	**(95% CI)**	**p** [f]
Enterococcus	Gastrointestinal illness	30 / 4251	1.17	(0.49 , 2.80)	1.11	(0.46 , 2.69)	10 / 1297	2.90	(0.75 , 11.16)	0.29	1.81	(0.46 , 7.09)	0.56
> 35 CFU [a]	Diarrhea	24 / 4285	1.59	(0.64 , 3.96)	1.61	(0.64 , 4.04)	9 / 1305	4.39	(0.92 , 20.98)	0.30	2.97	(0.67 , 13.21)	0.52
	Sinus pain or infection	44 / 4130	1.52	(0.70 , 3.31)	0.99	(0.38 , 2.54)	19 / 1262	2.16	(0.85 , 5.48)	0.57	1.87	(0.73 , 4.81)	0.36
	Earache or infection	38 / 4233	1.28	(0.59 , 2.80)	1.26	(0.57 , 2.77)	14 / 1274	1.26	(0.42 , 3.79)	0.98	1.30	(0.44 , 3.81)	0.96
	Infection of open wound	9 / 4360	1.14	(0.22 , 5.87)	1.26	(0.23 , 6.91)	6 / 1332	1.28	(0.26 , 6.26)	0.92	1.90	(0.39 , 9.23)	0.74
	Skin rash	19 / 4230	2.38	(0.90 , 6.34)	1.77	(0.54 , 5.85)	5 / 1267	0.82	(0.14 , 4.89)	0.30	0.38	(0.04 , 3.76)	0.22
	Fever	22 / 4366	1.48	(0.55 , 4.01)	1.50	(0.57 , 3.94)	2 / 1342	-- [g]			-- [g]		
	Upper respiratory illness [c]	37 / 3679	0.64	(0.26 , 1.60)	0.48	(0.19 , 1.19)	15 / 1090	1.33	(0.48 , 3.72)	0.31	1.37	(0.52 , 3.60)	0.13
	Any infectious symptom [d]	50 / 4080	1.08	(0.54 , 2.17)	1.01	(0.47 , 2.15)	17 / 1264	3.08	(1.10 , 8.56)	0.11	2.76	(0.94 , 8.16)	0.14
Enterococcus	Gastrointestinal illness	30 / 4251	0.93	(0.38 , 2.25)	0.92	(0.38 , 2.22)	10 / 1297	2.69	(0.69 , 10.39)	0.23	1.78	(0.44 , 7.26)	0.46
> 104 CFU [b]	Diarrhea	24 / 4285	1.26	(0.50 , 3.18)	1.32	(0.52 , 3.34)	9 / 1305	4.06	(0.85 , 19.50)	0.25	2.90	(0.59 , 14.21)	0.44
	Sinus pain or infection	44 / 4130	1.85	(0.94 , 3.63)	1.29	(0.58 , 2.85)	19 / 1262	1.63	(0.65 , 4.10)	0.83	1.33	(0.50 , 3.53)	0.96
	Earache or infection	38 / 4233	1.29	(0.68 , 2.47)	1.25	(0.63 , 2.49)	14 / 1274	1.14	(0.38 , 3.43)	0.84	1.11	(0.40 , 3.11)	0.85
	Infection of open wound	9 / 4360	2.53	(0.58 , 11.02)	2.72	(0.60 , 12.28)	6 / 1332	2.32	(0.43 , 12.47)	0.94	3.00	(0.54 , 16.65)	0.93
	Skin rash	19 / 4230	1.47	(0.56 , 3.91)	1.03	(0.33 , 3.16)	5 / 1267	0.74	(0.13 , 4.37)	0.51	0.27	(0.03 , 2.71)	0.28
	Fever	22 / 4366	1.44	(0.55 , 3.79)	1.45	(0.55 , 3.83)	2 / 1342	-- [g]			-- [g]		
	Upper respiratory illness [c]	37 / 3679	0.92	(0.43 , 1.96)	0.71	(0.34 , 1.51)	15 / 1090	1.22	(0.44 , 3.39)	0.66	1.08	(0.40 , 2.90)	0.51
	Any infectious symptom [d]	50 / 4080	1.56	(0.83 , 2.94)	1.52	(0.81 , 2.86)	17 / 1264	2.79	(0.99 , 7.82)	0.44	2.51	(0.84 , 7.47)	0.45

CFU: colony forming units

[a] Daily geometric mean > 35 CFU at any point during three-day follow-up

[b] Any single sample > 104 CFU at any point during three-day follow-up

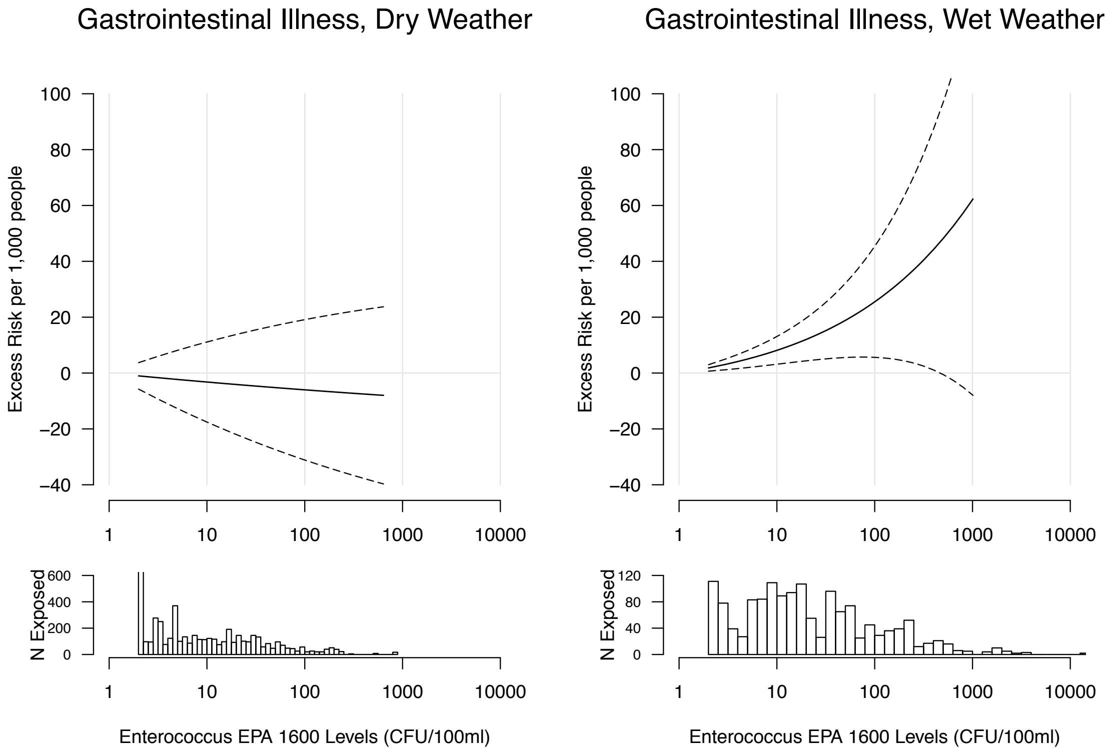
[c] Only measured in year 2

[d] Includes gastrointestinal illness, diarrhea, vomiting, eye infections, infected cuts and fever

[e] Adjusted for a range of covariates (see statistical methods for details)

[f] p-value for interaction term between water quality indicator and dry vs. wet weather

[g] Could not estimate due to sparse data



Web Figure 6: Excess risk of gastrointestinal illness associated with *Enterococcus* levels measured during dry and wet weather periods, predicted from a log-linear model among surfers at Tourmaline Surfing Park and Ocean Beach, San Diego, CA during the winters of 2013-14 and 2014-15. Dashed lines indicate model-based 95% confidence intervals and histograms show the distribution of *Enterococcus* exposure in the population during dry and wet weather periods.

Web Table 5: Negative control exposure analysis that matched daily mean log10 fecal indicator bacteria levels to surfers who were enrolled on that day, but had no seawater exposure for the past 5 days. Adjusted incidence rate ratios (IRRs) estimate the association between a 1-log10 increase in fecal indicator bacteria levels and incidence illness, stratified by dry and wet weather periods.

		Dry weather		Wet weather		
		IRR [c]	**(95% CI)**	**IRR [c]**	**(95% CI)**	**p [e]**
***Enterococcus* $\mathbf{log_{10}}$**	Gastrointestinal illness	1.69	(0.80 , 3.56)	1.13	(0.39 , 3.25)	0.56
	Diarrhea	1.43	(0.68 , 3.04)	0.82	(0.26 , 2.64)	0.46
	Sinus pain or infection	1.27	(0.64 , 2.50)	1.24	(0.35 , 4.42)	0.97
	Earache or infection	2.07	(0.60 , 7.10)	1.17	(0.31 , 4.32)	0.54
	Infection of open wound	-- [d]		-- [d]		
	Skin rash	0.88	(0.25 , 3.13)	-- [d]		
	Fever	1.21	(0.56 , 2.62)	0.82	(0.14 , 4.66)	0.69
	Upper respiratory illness [a]	1.34	(0.73 , 2.46)	0.46	(0.12 , 1.68)	0.15
	Any infectious symptom [b]	1.62	(0.86 , 3.04)	0.97	(0.33 , 2.84)	0.43
Fecal coliform $\mathbf{log_{10}}$	Gastrointestinal illness	1.69	(0.54 , 5.31)	0.46	(0.09 , 2.34)	0.22
	Diarrhea	1.19	(0.32 , 4.36)	0.64	(0.13 , 3.07)	0.58
	Sinus pain or infection	1.08	(0.39 , 2.98)	0.72	(0.16 , 3.29)	0.67
	Earache or infection	2.43	(0.47 , 12.66)	1.77	(0.32 , 9.83)	0.79
	Infection of open wound	-- [d]		-- [d]		
	Skin rash	0.35	(0.05 , 2.75)	-- [d]		
	Fever	0.75	(0.17 , 3.25)	0.32	(0.06 , 1.67)	0.44
	Upper respiratory illness [a]	0.78	(0.28 , 2.20)	0.34	(0.10 , 1.13)	0.33
	Any infectious symptom [b]	1.29	(0.47 , 3.53)	0.34	(0.08 , 1.54)	0.17
Total coliform $\mathbf{log_{10}}$	Gastrointestinal illness	1.15	(0.49 , 2.70)	0.84	(0.22 , 3.16)	0.70
	Diarrhea	0.81	(0.31 , 2.11)	0.80	(0.21 , 3.11)	0.99
	Sinus pain or infection	1.35	(0.71 , 2.55)	1.28	(0.31 , 5.25)	0.95
	Earache or infection	2.64	(1.17 , 5.97)	0.76	(0.32 , 1.79)	0.04
	Infection of open wound	-- [d]		-- [d]		
	Skin rash	0.96	(0.22 , 4.17)	-- [d]		
	Fever	1.61	(0.74 , 3.51)	0.90	(0.11 , 7.24)	0.61
	Upper respiratory illness [a]	1.69	(0.88 , 3.21)	0.47	(0.14 , 1.53)	0.06
	Any infectious symptom [b]	1.14	(0.57 , 2.28)	1.02	(0.28 , 3.70)	0.88

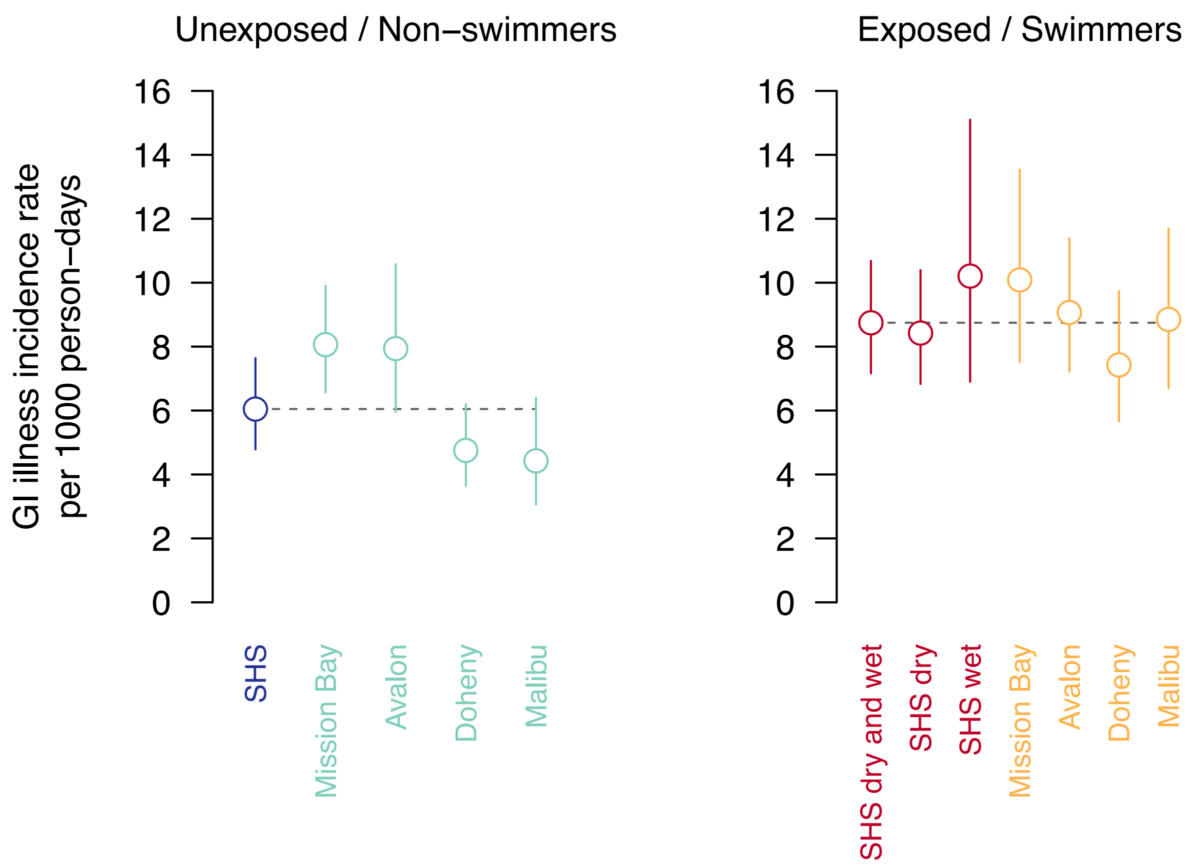
[a] Only measured in year 2

[b] Includes gastrointestinal illness, diarrhea, vomiting, eye infections, infected cuts and fever

[c] Adjusted for a range of covariates (see statistical methods for details)

[d] Could not estimate due to sparse data

[e] p-value for interaction term between water quality indicator and dry vs. wet weather



Web Figure 7: Incidence rates of gastrointestinal (GI) illness per 1,000 person-days in the present study (denoted SHS) and four previous summer swimmer cohort studies conducted in California. Swimmer cohorts were limited to adults (18 years or older) and swimmers included those with head immersion exposure. Rates in the present study are presented overall (dry and wet weather periods) and stratified by dry and wet weather. Vertical lines mark 95% confidence intervals. Horizontal dashed lines mark the present study rates to facilitate comparison with other estimates. Mission Bay = Mission Bay, San Diego [1]; Avalon = Avalon beach, Catalina island [2]; Doheny = Doheny State Beach [3]; Malibu = Malibu Surfrider State Beach [4].

References

1. Colford JM, Wade TJ, Schiff KC, et al. Water quality indicators and the risk of illness at beaches with nonpoint sources of fecal contamination. Epidemiology 2007;18:27–35.

2. Yau VM, Schiff KC, Arnold BF, et al. Effect of submarine groundwater discharge on bacterial indicators and swimmer health at Avalon Beach, CA, USA. Water Res. 2014;59:23–36.

3. Colford JM, Schiff KC, Griffith JF, et al. Using rapid indicators for Enterococcus to assess the risk of illness after exposure to urban runoff contaminated marine water. Water Res. 2012;46:2176–2186.

4. Arnold BF, Schiff KC, Griffith JF, et al. Swimmer illness associated with marine water exposure and water quality indicators: impact of widely used assumptions. Epidemiology 2013;24:845–853.